ACADEMIC SELF-CONCEPT, TEACHER QUALITY AND PEER EFFECTS ON CHINESE SECONDARY SCHOOL STUDENT PERFORMANCE AND TRACK CHOICE

A Dissertation

by

YAOJING WANG

Submitted to the Office of Graduate and Professional Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Chair of Committee,	Li Gan
Co-Chair of Committee,	Mark Hoekstra
Committee Members,	Timothy J. Gronberg
	Lori L. Taylor
Head of Department,	Timothy J. Gronberg

May 2016

Major Subject: Economics

ABSTRACT

In this dissertation, we use student-level high school and college entrance exam data from a city in China to explore the effects of peers, relative ranking and teacher quality on student test outcomes and track choices.

The broad consensus within the literature of peer effects suggests that students are positively impacted by higher-quality peers. A potential competing theory proposes that students who are near the top of the distribution of their schools—"big fish in a small pond"—may actually benefit more relative to students of similar ability who attend better schools. We estimate the returns to being a "big fish in a small pond," in addition to traditional measures of peer effects. Results provide evidence for the existence of both effects.

Besides peer effects, the role of school quality is also an important factor to student outcomes. We use a regression discontinuity design that compares applicants barely above and below high school admission thresholds. Results show significant academic gains from attending elite Tier I high schools. Further evidence suggests that these returns to high school quality are driven by teacher quality, rather than by peer quality or class size.

Having high-quality peers not just potentially improves student test score outcomes; it could also influence their choice in major. We use the fixed effect model to compare within school variations across adjacent cohorts and find that high school girls

are more likely to choose the science track if there is a higher percentage of high performing girls in her class. High performing boys have less or potentially negative effect on girls' tendency of choosing science. Boys, on the other hand, are not affected by their peers in track choice.

For Luke,

without whom this dissertation would have been completed one year earlier.

ACKNOWLEDGEMENTS

I would like to thank my committee chair, Dr. Li Gan, for guidance and patience with my research and all the paperwork. Thank you for tolerating my slow progress and never realize the threat to throw me out on the street. Thank you to my co-chair, Dr. Mark Hoekstra, for continuous encouragement and guidance, all the Wednesday 1pm phone calls and the awesome bib. Thank you also to my committee members, Dr. Timothy Gronberg and Dr. Lori Taylor for their support and help throughout the course my graduate work at A&M.

Thank you to Courtney, for spending numerous nights in front of Stata in my office and drawing monsters on the white board every time she visited. Thank you to Pierre, for being the more reliable part of the team and ending every task assignment with “take your time, no hurry”.

Thank you also to all my friends, classmates, professors and secretaries in the department of Economics, who made this extended journey as enjoyable as it could be.

Thanks to my PKU professor Hu Tao who taught me Economics 101. He really cared about his students and showed real passion for Economic research.

Finally, thank you to my parents for always believing in me and offering me constant emotional and financial support. Special thanks to my husband and son, for saving me in so many ways.

TABLE OF CONTENTS

Page

LIST OF FIGURES

LIST OF TABLES

CHAPTER I

INTRODUCTION

Students faced with the decision of where to attend school consider many factors—teaching quality, programs offered, reputation, and location—just to name a few. These factors are relevant for students making the college decision, as well as for students (and parents) who have multiple available options for primary and secondary school. Students at all levels of schooling may also consider the quality of the peers at these potential schools and how their own ability compares with their future classmates.

At the end of the day, students care about outcomes. These outcomes may be test score outcomes, college outcomes and labor market outcomes. In the context of out data, test score is the one and only determinants for high school and college admissions. Most high schools and the students have one common goal, that is to maximize the test score in the College Entrance Test. Apparently high school choices are highly endogenous. In fact, there is strong sorting into high schools. Here in this dissertation, we apply three methods to solve the endogeneity issue in school choice.

In the first, we attempt to evaluate the "big fish small pond" theory in psychology. Most Chinese students have strong preferences towards quality schools. Some of them are even willing to pay a high price to get into a better school. However, being at the bottom of a better school can impose a negative psychological effect, which might negate the initial purpose of going to a better school. To address the endogeneity issue in school choice, we use a two-stage-least-square method, instrumenting the endogenous peer quality with predicted school quality, to estimate the effects of peer mean scores and within-cohort ranking on test score outcomes. Being with better peers and being a "big fish" in the pond both help improve test outcomes but we also find evidence of the negative impact of being a "small fish" on student performances.

In the second paper, we take advantage of the high school admission cutoffs and apply the regression discontinuity model to compare the test score results of those students who are just above or below the cutoffs. Since the students take the High School Entrance Test before the cutoff scores are generated and the admission process is centralized and computerized, there is virtually no chance for a student to manipulate her scores around the cutoff. Whether a student is just above or below a cutoff is essentially random. Data shows that peer quality jumps by around 0.2 standard deviations once a student crosses a threshold. However, test scores only increases significantly if a student passes the cutoff for an elite school. What differentiate an elite school from the rest seem to be the school quality, specifically, teacher quality. As we examine through all measureable factors such as class size, we find that the only factor that distinguish the elite schools is the concentration of superior teachers. In China, teachers are ranked by their tenure, performance and various attributes and our results support that these superior teachers do seem to improve student performance.

In the third paper, we explore the peer effect on student track choice. In China, students choose the science or arts track as early as high school. The literature also shows that females' interest in science and engineering starts to divert from males' since high school. With our unique data, we are able to look into the question of why there are fewer women the science fields at the high school stage. I use the fixed effects model to compare within school variations in science track choices across adjacent cohorts. This way I eliminate school effects which can be endogenous and a threat to identification. Evidence displays a heterogeneous treatment effect of high performing peers on girls and boys where boys are hardly affected by their peers about choosing the science track. For girls, having more girl classmates who are good at math can increase their likelihood of choosing science while boy classmates have a smaller to negative impact.

CHAPTER II

DOES IT PAY TO BE A BIG FISH IN A SMALL POND? EVIDENCE FROM CHINESE SCHOOLING DATA

II.1 Introduction

There is a large body of literature examining the impact of a student's peers and classmates on his own actions and outcomes. The general consensus within the peer effects literature is that there is that individuals are influenced by the behavior of the people around them, and there is substantial evidence that this relationship is causal, not simply correlative. Sacerdote (2001) identifies the existence of peer effects at the college level by exploiting the random assignment of college roommates. He finds evidence that having a roommate with a higher GPA positively affects an individual's own GPA and that living in a dorm with members of a particular fraternity increases an individual's probability of joining that same fraternity. Several other papers examine the existence of peer effects at the secondary school level. Ding and Lehrer (2007) use schooling data from a province in China where students are assigned to schools based only on observable characteristics. Their results show that peer effects exist but that they are different for different types of students. They find the strongest evidence of effects for high scoring students, suggesting that these students benefit more from high scoring peers than do their lower scoring counterparts. Lavy et al. (2011) also examine effects at the secondary school level. Using administrative data including students' birthdays to identify students who have repeated a grade and are likely lower performing students, they examine the impact of having a high proportion of low achievers within a class. They find that a higher proportion of these students has a negative impact on the performance of their classmates.[1]

The conclusions of the peer effects literature would lead a student to believe that, given the choice, he should attend the school with the highest peer quality, all else equal. If students are positively influenced by the performance of their peers, then a student who is at the bottom of the score distribution of his school will benefit more than he would were he at the top of the distribution at a poorer quality school. A competing theory, however, proposes that the effect of a student's relative ranking may move in the opposite direction. A student who is a "big fish in a small pond" may be in a lower quality school overall, but he will be the best among his peers. It is possible that this student may benefit *more* than he would at a higher quality school where he would be a "small fish in a big pond." This effect could work through several channels. There could be a positive psychological effect for an individual who performs better than the majority of his peers. It may also be the case that more resources or additional attention from teachers or other mentors could enhance his educational experience and increase his future achievement levels.

This effect is referenced by Marsh (1984) within the education literature in an analysis of ability grouping, its impact on self-concept, and the subsequent effects on student outcomes. Because students consider themselves in reference to their peer

---

[1] A related strain of the literature examines schools' tracking policies and their impacts on student performance. See Figlio and Page (2002), Lefgren (2004), Duflo et al. (2011), and Collins and Gan (2013).

groups, including their classmates, the effect—which he calls the "big-fish-little-pond-effect" could have a positive impact on students' self-concepts. Seaton, Marsh, and Craven (2010), using PISA (Program for International Student Assessment) test score data and survey data, show that the big fish effect is present across a variety of different cultures and student characteristics[2].

In this paper, we attempt to analyze the impact of being a big fish in a small pond on academic performance as measured by test scores. We employ a unique student-level data set of Chinese high school students that includes high school entrance and college entrance scores, as well as information on how students are assigned to high schools. Using a school entrance model based on specific schools' test cutoff scores, we identify both "big fish" and "small fish" and analyze how their relative positions within the score distributions of their schools impacts their performance on the college entrance exam. The results suggest that school quality has a positive impact on score, but that after controlling for both school quality and student quality, there is still a separate big fish effect. The big fish effect suggests that some students may benefit from "dropping down" from their best possible school choice and attending a school in which they place higher within the distribution of their peers.

---

[2] See also Davis (1966) for an early explanation of the big fish effect within the sociology literature.

II.2 Model

II.2.1 A Simple Two School Entrance Model

Consider the following basic model of test scores:

$$s_{idt} = \alpha + \rho s_{idt-1} + X_{idt}\beta + u_{idt}, \tag{II.1}$$

where score *s* of student *i* from district *d* in time *t* is a function of his score in the previous period and a vector *X* of student-specific characteristics. Students are placed into secondary schools based only on the entrance exams they took in the previous period, measured here by s_{idt-1} . For simplicity, assume that there are only two schools and that school 1 is of higher quality than school 2. There are two cutoff scores for the entrance exam, c_1 and c_2, used to determine which school students will attend. If a student's score is above the first cutoff, c_1, he is automatically eligible to attend school 1, but he has the option to "drop down" to school 2, if he chooses. School 1 may also choose to accept some students whose scores are below the first cutoff c_1 but above the second cutoff c_2. These students must pay an additional fee in order to enroll in school 1. Students whose scores are below the second cutoff c_2 do not have an option and must enroll in school 2. Therefore, each student *i* falls into one of three categories:

<u>High Scorers</u>:	If $s_{idt-1} > c_1$,	student *i* may attend either school 1 or school 2.
<u>Marginal Scorers</u>:	If $c_2 < s_{idt-1} < c_1$,	student *i* may attend school 2 or pay an additional fee to attend school 1.
<u>Low Scorers</u>:	If $s_{itd-1} < c_2$,	student *i* must attend school 2.

In the model above, most high scoring students will attend school 1, and most low scoring students will attend school 2. However, some high scoring students who qualify for school 1 will actually enroll in school 2. These students are categorized as big fish in a small pond. Their scores make them eligible to attend a higher quality school, but they enroll in a lower quality school. Relative to the other students in their "pond," they are the "biggest fish." We can also identify small fish in a big pond using the cutoff points given above. Students whose scores fall between the two cutoff points are only eligible to attend school 2, unless they pay to enroll in the higher quality school. These students will end up being the lowest-scoring students at a high quality school, making them small fish in a big pond.

Using the cutoff scores, the students' actual scores, and the enrollment records for the schools, we create the following two dummy variable groups:

big fish in a small pond: $BF_i = 1$ if $s_{idt-1} > c_1$ and student i enrolls in school 2

small fish in a big pond: $SF_i = 1$ if $c_2 < s_{itd-1} < c_1$ and student i enrolls in school 1

We estimate the effects of both groups by adding the dummy variables to the regression in equation (II.1).

$$s_{idt} = \alpha + \rho s_{idt-1} + \gamma_1 BF_{idt} + \gamma_2 SF_{idt} + X_{idt}\beta + u_{idt}. \tag{II.2}$$

If γ_1 is positive, then students who qualify for a higher quality school but "drop down" to a lower quality school score better, on average, than students who do not. In other words, there is an added benefit to being a big fish in a small pond. If the big fish in a small pond story holds, we would also expect that being a small fish in a big pond would result in lower than expected scores. This would mean that γ_2 would be negative. In this case, students whose scores qualify them for a lower quality school but who pay to attend a higher quality school do worse, on average, than students who do not.

A typical peer effects story should yield the opposite results for the coefficients. If the quality of a student's peers has a significant direct impact on his own performance, we would expect that γ_1 is negative, or that big fish students actually perform worse than average, because they are now surrounded by lower achieving peers. Similarly, we would expect that γ_2 is positive, or that small fish students perform better than average, because they are now surrounded by higher achieving peers.

II.2.2 A School Entrance Model with Many Potential Schools

The model explained here involves only two schools, but expanding it to include many schools of differing qualities requires only the addition of school-specific cutoffs. We now define two cutoffs for each school j, c_{1j} and c_{2j}, analogous to the cutoffs described above. Each student now falls into one of three categories for each school:

<u>High Scorers</u>:	If $s_{idt-1} > c_{1j}$,	student i may attend school j without an additional fee.

Marginal Scorers: If $c_{2j} < s_{idt-1} < c_{1j}$, student i may attend school j if he pays an additional fee.

Low Scorers: If $s_{idt-1} < c_{2j}$, student i is ineligible to attend school j.

The dummy variables for big fish and small fish are similar to the ones in the two school model. We define the following two groups:

big fish in a small pond: $BF_{idt} = 1$ if $s_{idt-1} > c_{1j}$ and student i enrolls in a school of lower quality than school j

small fish in a big pond: $SF_{idt} = 1$ if $c_{2j} < s_{idt-1} < c_{1j}$ and student i enrolls in school j

The regression, which allows for identification of the big fish effect, is comparable to equation (II.2), except we have added vector C_{idt}, which includes school-specific, or campus-specific controls.

$$s_{idt} = \alpha + \rho s_{idt-1} + \gamma_1 BF_{idt} + \gamma_2 SF_{idt} + X_{idt}\beta + C_{idt}\theta + u_{it}. \tag{II.3}$$

As stated earlier, positive values for γ_1 and negative values for γ_2 would support the big fish in a small pond story. Negative values for values for γ_1 and positive values for γ_2 would provide evidence for peer effects.

II.3 The High School Admissions Process and Data Description

In China, admission to high school is based mainly on students' performance on a high school entrance test (HET) administered at the city-level. At the end of middle school (ninth grade), students indicate their high school preferences. They are then tested in several subjects—Chinese, English, mathematics, physics, chemistry, politics, and physical education. Students' total scores from the HET are calculated and used in the high school admissions process[3]. High schools predetermine a cohort size of x incoming students and admit the highest-scoring x students who have indicated the school as a preference. The lowest HET score among the admitted students' scores is labeled as the cutoff score of the high school and is announced publically.

An additional feature of the high school admission process is the existence of "high priced seats." In order to raise additional funds for high schools, the government allows schools to designate 10% of enrollment as "high priced seats"—paid enrollment slots available to students who may not surpass the regular cutoff of the school. The school creates an additional high price cutoff, which is lower than the regular cutoff, for students wishing to obtain admission under this feature. Students scoring above the regular cutoff may be admitted to the school at no cost, and students who score above the high price cutoff but below the regular cutoff may pay the high price in order to obtain admission. Students scoring below the high price cutoff are ineligible for

---

[3] The maximum total HET score is 790. The Chinese, English, and mathematics sections each account for 150 possible points. The physics, chemistry, and politics sections each account for 100 possible points. The physical education section accounts for 40 possible points.

admission[4]. When students apply for high school admission at the end of middle school, they are required to specify if they wish to apply for regular admission or for the high priced seats. Once enrolled, the student pays a one-time fee, which varies by school, but is usually around 40,000 RMB (about $6,500). After this fee is paid, there are no additional differences in registration or tuition for these students.

The admission process begins after the HET scores are released, which is usually two to three weeks after the students take the exam. The process is coordinated by the city's admission center officials who have all students' scores, students' submitted preference lists, and the entering class size for each school. Schools are divided into four tiers, with the best schools in tier 1. The admission process begins with tier 1 schools; tier 2 schools will not enter the process until all seats have been filled in tier 1 schools. Within the tiers, schools are ranked by quality, with school A being the highest quality. All students who have indicated school A as their first-choice preference are ranked by their HET scores in descending order. If school A plans on an entering class size of x, the top x students on the list are admitted to school A. If the list of students who have indicated school A as their first-choice preference is smaller than x, then the remaining seats will be filled by students who have indicated school A as their second choice. Next, the seats in school B are filled by an analogous process. This continues until all the seats in all the tiers have been filled.

---

4 One exception to this is students labeled as having "special skills," such as unusual athletic or artistic ability. Some of these students may be admitted to a school even though their scores are below both cutoffs, but these instances are rare.

The college admissions process works in a similar manner. After three years in high school, college-bound students take a college entrance test (CET) administered at the province-level. A student's score on the CET is the main determinant (and in most cases the only determinant) of his admission into college.

Table II.1 High School Summary Statistics

	obs	mean	std dev	min	max
high school size	49674	553.3024	324.2068	1	1531
female	49674	0.5209768	0.4995648	0	1
high school entry score	47267	617.1739	62.2691	280	759
college entry score	49674	501.4739	98.23079	99	709
cutoff (free)	48114	584.4709	70.38671	380	711
cutoff (price)	45840	547.7818	85.63518	380	685
paid high price	49674	0.0799211	0.2711737	0	1
different district	49674	0.1850264	0.3883229	0	1
suburb	49674	0.4291581	0.494961	0	1
rural	49674	0.1533599	0.3603377	0	1

We utilize a unique student-level data set of high school and college entrance test scores for four cohorts (each including about 48,000 students) from 135 different high schools in a large metropolitan area in China with a population of over ten million. The data was obtained from the city's testing center. The four cohorts of students began high school in the years 2004-2007, when we observe their incoming HET scores, and completed high school in 2007-2010, when we observe their outgoing CET scores[5]. In

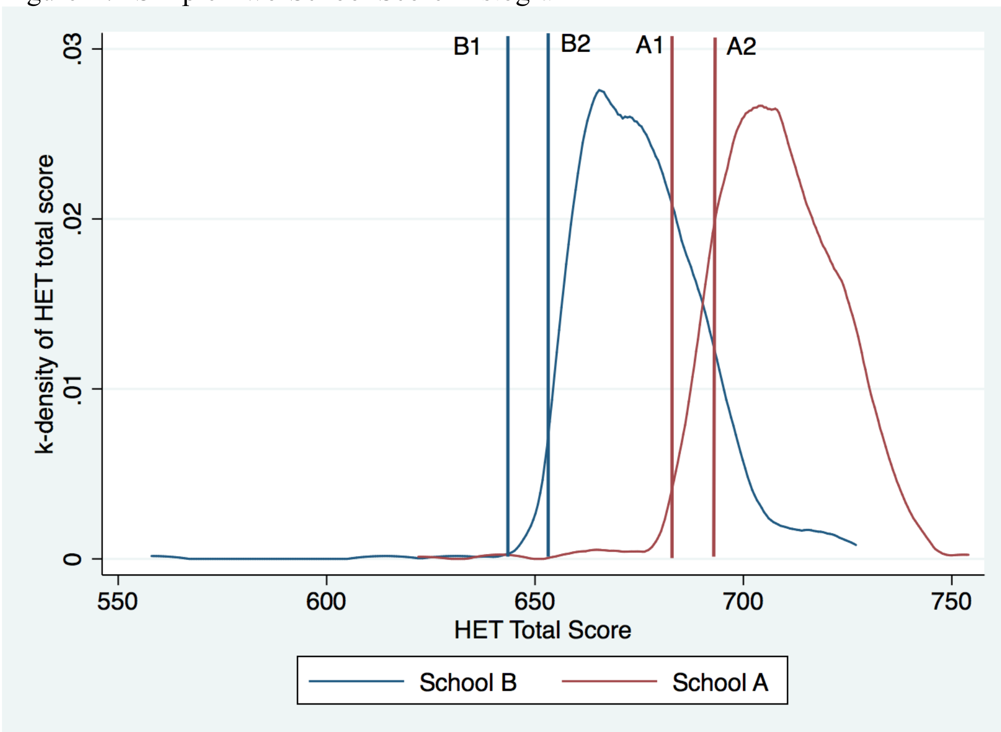
[5] Almost all of the students complete high school in three years. It is rare in China for a student to repeat a grade. A few students choose to retake the CET the year following graduation. However, we only observe the students' first attempts (as immediate high school graduates).

addition to this score information, we observe basic demographic characteristics of students, such as gender, minority status[6], date of birth, parents' occupation, and middle school district. The data set also includes the cutoff scores for each high school for both regular and high priced seats. Summary statistics are shown in table II.1.

We use the cutoff scores for each high school to categorize students as big fish and small fish. Students whose score above the cutoff for a particular high school (and would therefore qualify for admission to that school) but who attend a school with a lower cutoff are tagged as big fish. Figure II.1 shows a simple example of student score histograms of two high schools of differing quality. School A, the higher quality school, has a regular cutoff of A2 (HET score=693) and a high price cutoff of A1 (HET score=685). School B, the lower quality school, has a regular cutoff of B2 (HET score=659) and a high price cutoff of B1 (HET score=653). Students who score above A2 but attend school B are categorized as big fish. These students' scores qualified them for admission to a higher quality school, but they ended up attending a lower quality one. The farther these students' scores are from the A2 cutoff, the stronger the effect should become.

[6] The city is predominantly Han; less than one percent of students are considered ethnic minorities.

Figure II.1 Simple Two-School Score Histogram



Similarly, we categorize a student as a small fish if his score should assign him to the lower quality school (school B), but he attends school A. There are two groups of students who fall into this category. Students who score between A1 and A2 are eligible to pay for high priced seats in school A. There are also some students who score below A1 who still obtain admission to school A, contrary to the typical admissions policy[7].

Figure II.1 indicates that for students categorized as big fish, there is substantial variation in the difference between a student's score and the cutoff of the school he

---

[7] These students are almost certainly not randomly assigned and are likely to be systematically different from similar scoring students who attend school B. The estimation of this effect should be viewed as descriptive rather than causal because of the likely bias caused by this type of selection.

actually attends, suggesting that there is variation in "how big of a fish" a student actually is, or how far he drops down from his best possible school. Students in the far right tail of the score distribution of school B scored high enough not only to gain admission to school A, but also to any other higher quality schools whose cutoffs fall between A2 and the score of the student. If a big fish effect exists, it should be greater for these students than for the students whose scores are just slightly above the A2 cutoff.

To test this, we create a variable indicating how far a student drops down (or jumps up) from his best possible school. In order to determine this, we first define a choice set of schools for each student, given his high school entrance score s_{idt-1}. Intuitively, we define a school as being in a student's choice set if he is eligible to enroll in that school without paying an additional fee. The following two qualifications must be met in order for school *j* to be a member of the choice set of student *i*.

(i) Student *i*'s score must qualify him for entrance to school *j*. That is, $s_{idt-1} \geq c_{1j}$.

(ii) School *j* must accept students from district *d*.

Because not all schools accept students from every district, condition 2 could exclude some students from a particular school even if the students' scores are above the school cutoff. (Additionally, some schools have different cutoffs for students from different districts. We incorporate these differences into the schools' cutoff rules.)

Given the choice set of student i, we define his best possible school as the highest ranked school he would be eligible to attend without paying an additional fee. Schools are ranked according to their cutoff scores. The cutoff of the best possible school is given by the following:

$$c_1(best_{idt}) = max\{\text{all } c_1 \text{ cutoffs of schools in choice set of student } i\} \quad \text{(II.4)}$$

After determining the cutoff of the best possible school for student i, we can calculate how far he dropped down or jumped up, relative to the best school into which he could have enrolled without paying an additional fee. Using the cutoff score for the school in which student i actually enrolls, we calculate the following variable:

$$fish\ size_{idt} = c_1(best_{idt}) - c_1(actual_{idt}). \quad \text{(II.5)}$$

This measure shows the difference between cutoffs of the highest ranked school for which student i qualifies and the school he actually chooses to attend. For each individual student, the following relationships should hold. For students who are big fish in a small pond, $BF_{idt} = 1$, and $fish\ size_{idt} > 0$. Students attend schools with lower rankings relative to the best schools they can attend. As $fish\ size_{idt}$ becomes more positive, students drop down to schools with even lower rankings, relative to their best possible options. That is, those students become "bigger fish" relative to their classmates.

For students who are small fish in a big pond, $SF_{idt} = 1$, and $fish\ size_{idt} < 0$. Students attend schools with higher rankings relative to the best schools they are eligible to attend. Many of these students pay to receive a seat in a higher ranked school because their scores fall between the two cutoffs for the school. (In reality, some students who jump up to better schools may not actually be tagged as paying the additional fee. Some of these students may obtain entry because of special skills, such as athletic or fine arts abilities.) As $fish\ size_{idt}$ becomes more negative, students jump up to schools with even higher rankings, relative to their best no-fee options.

Table II.2 summarizes this distance measure for small fish and big fish, in addition to students who attend their best possible school. About 28 percent of the students in the data set gain admission to the best school for which their scores qualify them. Because they attend their best possible school, the actual cutoff of their school is equal to the best possible cutoff, and the fish size variable is equal to zero. The majority of students in the data set—about 64 percent—drop down and are categorized as big fish. There could be several reasons a student could attend a school that is of lower quality than his best qualifying school. The first is that geography could play a factor in his decision process and that he would rather attend a relatively lower ranked school closer to home than a relatively higher ranked one that is farther from home. The second reason could be that students list their school preferences before taking the high school entrance test, so uncertainty exists regarding what schools they should target[8]. The third reason is that students intentionally choose to drop down to an easier school. The implications of this choice on the estimation of the big fish effect are discussed in the results section.

Table II.2 Student Enrollments Relative to Student Scores

			fish size: best cutoff - actual cutoff			
	number of students	percent of total	mean	std dev	min	max
attend best possible school	12992	0.28	0	0	0	0
"drop down" to lower school	29316	0.64	37.311	38.26	1	293
"jump up" to higher school	3708	0.08	-44.932	37.59	-233	-1

Note: Fish size is the difference between the cutoff score of the best possible school minus the cutoff score for the school the student actually attends. Higher numbers indicate that the student has enrolled in a worse school relative to the school for which he qualified.

Because so many students in the data set are categorized as big fish, it is important to examine the distance between the two score cutoffs, shown in table II.2. The average cutoff difference is 37 points, and the greatest difference is 293 points. Although substantially fewer students are classified as small fish, we also summarize the difference for these students. Students in this group jump up an average of 45 points, and largest cutoff jump is 233 points.

Table II.3 breaks down the big fish and small fish groups by quintiles and shows the average rank of students in those categories relative to the rest of the students in their

[8] While the timing and uncertainty of the students' preference listing process does impact the high school outcome, a complete modeling of that decision process is outside the scope of the paper.

schools. Again, because the majority of the students in the data set drop down from their best possible school, a student who is tagged as a big fish does not necessarily rank in the highly in the distribution of scores within that school. Students in quintile 1 of the big fish category drop down only a little from their best possible school and on average only rank in the 33rd percentile of their actual school. Students in quintile 5 drop down a lot from their best possible school and on average rank in the 79th percentile of their school.

Table II.3 Percentile Rank of Students by Quintiles

		rank of student within current school			
big fish	obs	mean	std dev	min	max
quintile 1	5562	0.333	0.189	0.007	1
quintile 2	6131	0.537	0.199	0.014	1
qunitile 3	4798	0.610	0.185	0.031	1
quintile 4	6838	0.736	0.177	0.047	1
quintile 5	6002	0.785	0.227	0.002	1
small fish	obs	mean	std dev	min	max
quintile 1	616	0.029	0.032	0.001	0.184
quintile 2	724	0.078	0.134	0.002	0.978
qunitile 3	705	0.106	0.118	0.001	0.998
quintile 4	743	0.126	0.114	0.003	0.706
quintile 5	723	0.106	0.140	0.002	0.958

Note: Students in quintile 5 within "big fish" enroll in a school ranked much lower than their best-eligible school. Students in quintile 1 within "small fish" enroll in a school ranked much higher than their best-eligible school.

The opposite is true for the breakdown of the students in the small fish category. Students in quintile 1 of the small fish jump up a lot from their best possible school and on average rank in the 3rd percentile. Students in quintile 5 of this group enroll in a school only slightly better than their best possible school but still only rank in the 11th percentile, on average.

II.4 Results

II.4.1 OLS Results

Using this data, we examine the effect of school quality on performance, in addition to the big fish effect. The baseline results are shown in table II.4. Column 1 reports the results from a basic version of equation (II.1). The dependent variable is college entrance score (CET), standardized to have a mean of zero and a standard deviation of one. We include controls for standardized high school entrance score (HET), race, gender, choice of program (science or humanities), private or public school, whether the student paid the high price fee to attend a better school, and whether the student attends a high school in a different district than his middle school.

Column 2 includes the base regression with additional controls, peer mean and peer variance, which reflect the composition of students within the school the student attends. The results indicate the presence of peer effects. Controlling for their own test scores, students benefit from attending a school with higher achieving peers on average. A one standard deviation increase in the scores of his peers increases a student's predicted score by .19 standard deviations. In addition, students seem to benefit from having less score variation within their cohorts; higher peer variance decreases predicted scores. These results are consistent with findings from Ding and Lehrer (2007).

Table II.4 OLS Baseline Results

VARIABLES	(1)	(2)	(3)
high school entry score	0.7780***	0.6425***	0.5984***
	(0.0027)	(0.0049)	(0.0054)
art student	-0.1158***	-0.1171***	-0.1138***
	(0.0056)	(0.0055)	(0.0055)
female	0.0539***	0.0549***	0.0524***
	(0.0079)	(0.0078)	(0.0078)
race	0.0006	0.0307	0.0026
	(0.0303)	(0.0298)	(0.0298)
private school	0.0408***	0.1229***	0.1131***
	(0.0154)	(0.0161)	(0.0160)
paid high price	0.0125	-0.0466***	0.0384***
	(0.0093)	(0.0094)	(0.0108)
different district	0.0244***	-0.0175**	-0.0121*
	(0.0070)	(0.0070)	(0.0070)
rural	-0.0172	0.0177	0.0346***
	(0.0111)	(0.0113)	(0.0113)
suburb	-0.1717***	-0.1189***	-0.0956***
	(0.0080)	(0.0081)	(0.0082)
rural*female	0.0868***	0.0947***	0.0972***
	(0.0151)	(0.0148)	(0.0148)
suburb*female	0.1126***	0.1123***	0.1117***
	(0.0108)	(0.0106)	(0.0106)
peer mean		0.1906***	0.2594***
		(0.0078)	(0.0085)
peer variance		-0.2293***	-0.1954***
		(0.0208)	(0.0208)
big fish			0.0646***
			(0.0058)
small fish			-0.1591***
			(0.0116)
constant	0.1208***	0.1609***	0.1239***
	(0.0306)	(0.0312)	(0.0315)
observations	47,267	47,266	47,266
R-squared	0.6899	0.6984	0.7009

Notes: Standard errors in parentheses; ***p<0.01, **p<0.05, *p<0.1; both high school and college scores are standardized. Dependent variable is College Entrance Test (CET) score.

Column 3 adds two additional dummy variables, "big fish" and "small fish." These variables correspond to the definitions given in section 2.2. A student is defined to be a big fish if his high school entrance score qualifies him to attend a higher ranked school, but he drops down and enrolls in a lower ranked one. A student is defined to be a small fish if he jumps up to a school of a higher ranking than the best one that he qualifies for based on his score. On average, being a big fish in a small pond—or enrolling in a low ranked school relative to ability—increases predicted score by about .06 standard deviations, even after controlling for high school entrance score. Being a small fish in a big pond—or enrolling in a high ranked school relative to ability—decreases predicted score by about .16 standard deviations on average. This would suggest that there might be some benefit to choosing a school other than the best possible ranked school and that there may be a cost associated with attending a school that is too highly ranked, relative to a student's initial ability.

At this point, it is important to consider why a student is classified as either a big fish or a small fish and whether this might cause possible endogeneity in the model. As stated earlier, there are several reasons a student might drop down to a school to a school of lower quality than his best-eligible school. These reasons may be related to geography, uncertainty in the preference listing process, or unobserved student-specific characteristics such as ambition or effort. This last potential reason is particularly troubling and would cause a clear bias in the big fish effect. For example, students who are lazy or who are willing to put in less effort may specifically choose to apply to schools substantially lower than the ones for which their scores will qualify them.

Because effort should have a positive effect on college entrance exam, this would tend to cause a downward bias in the big fish effect in a simple model. Under this scenario, the true effect should actually be *bigger* than estimated coefficient in the underspecified model.

The opposite story could cause an upward bias in the small fish effect. This would be true if students who are classified as small fish are especially ambitious and hardworking, which is plausible, given their willingness to attend a school despite failing to explicitly qualify academically. However, there may exist other confounding factors for this group, such as a student's status as an athlete or his political connections, which could be correlated with other variables affecting college entrance test score.

We also examine heterogeneity in the big fish and small fish effect by dividing each dummy group into quintiles based on how far the student drops down or jumps up relative to his best possible school. Table II.5 presents these results. As explained in the data section, because so many students are classified as big fish, there is significant variation the difference between a student's current school and his best-eligible school. If students truly benefit from being a big fish in a small pond, then we expect the effect to be largest for students who drop down the most from their best possible schools. As explained in the data section, students in quintile 1 of the big fish category have scores that will qualify them for schools that are only a little better than their current school. Students in quintile 5 have scores that will qualify them for schools that are much better than their current school. Column 2 of table II.5 allows for heterogeneity in the big fish effect by quintile. As expected, the effect increases as the distance a student drops down increases. The effect is not significant for the bottom two quintiles, which include the students whose scores are closest to their schools' cutoffs. The estimated benefit of being a big fish is 0.03, 0.09, and 0.21 standard deviations for quintiles 3, 4, and 5, respectively. This matches the intuitive story that the average effect should be driven by the students who drop down the most relative to the schools for which they qualify.

We also allow for heterogeneity in the small fish effect. Students in quintile 1 of the small fish category have scores that are farthest below their schools cutoffs, and students in quintile 5 have scores that are only just below the cutoff. As expected, we find negative effects that are much larger in magnitude for quintiles 1, 2, and 3 (between -0.23 and -0.26 standard deviations) and smaller effects (between -.11 and -.14 standard deviations) for quintiles 4 and 5.

Table II.5 Heterogeneity in the Big Fish/Small Fish Effect

VARIABLES	(1)	(2)
peer mean	0.2594***	0.3285***
	(0.0085)	(0.0092)
peer variance	-0.1954***	-0.1931***
	(0.0208)	(0.0208)
big fish	0.0646***	
	(0.0058)	
small fish	-0.1591***	
	(0.0116)	
(big fish)*(quintile 1)		0.0112
		(0.0090)
(big fish)*(quintile 2)		-0.0048
		(0.0083)
(big fish)*(quintile 3)		0.0286***
		(0.0093)
(big fish)*(quintile 4)		0.0945***
		(0.0083)

Table II.5 Continued

VARIABLES	(1)	(2)
(big fish)*(quintile 5)		0.2082***
		(0.0087)
(small fish)*(quintile 1)		-0.2618***
		(0.0216)
(small fish)*(quintile 2)		-0.2252***
		(0.0219)
(small fish)*(quintile 3)		-0.2539***
		(0.0213)
(small fish)*(quintile 4)		-0.1400***
		(0.0209)
(small fish)*(quintile 5)		-0.1125***
		(0.0195)
constant	0.1239***	0.1103***
	(0.0315)	(0.0315)
observations	47,266	47,266
R-squared	0.7009	0.7045

Notes: Standard errors in parentheses; ***p<0.01, **p<0.05, *p<0.1; both high school and college scores are standardized; both regressions include all other controls shown in previous table.

II.4.2 2SLS Results

One potential problem with measuring the impact of school quality on students' college entrance test scores is that school quality is likely endogenous because students are selecting into particular schools. We attempt to address this problem by instrumenting for a student's actual school by exploiting the variation in the high school cutoffs to predict where a student should go to school based on his high school entrance test score. We compare a student's test score to the cutoffs for all possible schools and assign his predicted school as the best school he is eligible to attend given his test score[9]. We then use the quality of the student's predicted school as an instrument for the quality of the student's actual school.

Table II.6 Big Fish/Small Fish Effects With Student Percent Rank

VARIABLES	(1)	(2)	(3)	(4)	(5)
high school entry score	0.5986***	0.4291***	0.5626***	0.5368***	0.5008***
	(0.0053)	(0.0077)	(0.0060)	(0.0064)	(0.0068)
female	0.0518***	0.0530***	0.0513***	0.0519***	0.0525***
	(0.0077)	(0.0077)	(0.0077)	(0.0077)	(0.0077)
private school	0.1086***	0.0983***	0.1046***	0.1039***	0.1012***
	(0.0160)	(0.0159)	(0.0160)	(0.0160)	(0.0159)
paid high price	0.0500***	0.0555***	0.0442***	0.0479***	0.0522***
	(0.0109)	(0.0108)	(0.0109)	(0.0109)	(0.0108)
rural*female	0.0973***	0.0972***	0.0973***	0.0974***	0.0978***
	(0.0148)	(0.0146)	(0.0147)	(0.0147)	(0.0147)
suburb*female	0.1125***	0.1077***	0.1121***	0.1101***	0.1097***
	(0.0106)	(0.0105)	(0.0106)	(0.0106)	(0.0105)
peer mean	0.2635***	0.4638***	0.3087***	0.3389***	0.3807***
	(0.0086)	(0.0108)	(0.0092)	(0.0095)	(0.0099)
peer variance	-0.1837***	-0.1208***	-0.1667***	-0.1577***	-0.1449***
	(0.0209)	(0.0208)	(0.0208)	(0.0208)	(0.0208)
big fish	0.0602***	-0.0061	0.0576***	0.0488***	0.0258***
	(0.0060)	(0.0063)	(0.0060)	(0.0060)	(0.0062)
small fish	-0.1872***	-0.1630***	-0.1858***	-0.1759***	-0.1846***
	(0.0122)	(0.0121)	(0.0130)	(0.0132)	(0.0126)
percent rank		0.4487***			
		(0.0149)			
top 5% of class			0.1766***		
			(0.0122)		
bottom 5% of class			-0.0426***		
			(0.0142)		

[9] It is possible to achieve a better prediction in the first stage by limiting the schools to those within a certain distance of his home or by assigning a probability to each potential school using a multinomial logit model. We may explore this option in future drafts.

Table II.6 Continued

VARIABLES	(1)	(2)	(3)	(4)	(5)
top 10% of class				0.1737***	
				(0.0094)	
bottom 10% of class				-0.0691***	
				(0.0113)	
top 20% of class					0.1676***
					(0.0075)
bottom 20% of class					-0.0846***
					(0.0083)
observations	47,266	47,266	47,266	47,266	47,266
R-squared	0.7012	0.7069	0.7025	0.7034	0.7047

Notes: Standard errors in parentheses; ***p<0.01, **p<0.05, *p<0.1; both high school and college scores are standardized

The first stage results are shown in table II.6. We use several measures of school quality and report the results for each measure in a separate column. The four measures are the current cohort's incoming HET scores, the average HET scores for all cohorts at the school when the incoming cohort arrives, the HET scores of the cohort one year older than the incoming class, and the CET scores of the students who graduate as the incoming cohort arrives. We calculate each of these measures for both the actual school a student attends and the school we predict he should attend given his score and the cutoff scores of the high schools. The variable *school-level HET (expected school)* is the measure for the school we predict the student should attend. Table II.6 shows that in all specifications, there is a significant relationship between the expected and the actual school.

After predicting school quality for a student's actual school, we plug this measure into the main equation, with a student's college entrance test (CET) as the dependent variable. This effect of school quality in this model is now based on exogenous variation. The results are shown in table II.7. After controlling flexibly for a student's score, school quality still has a positive and significant effect, although the magnitude of the results is somewhat smaller than the OLS, depending on the specification. We also include dummy variables for big fish and small fish in this regression and find qualitatively similar results to the OLS specifications. Even after controlling for a student's own score and the quality of his school, there is still an additional benefit to being a big fish in a small pond and an additional cost to being a small fish.

Table II.7 Effects of School Quality (Second Stage)

	Dependent variable: College Entrance Test (CET) score			
	(1)	(2)	(3)	(4)
VARIABLES	Current Incoming Scores	Average (Past 3 Years)	Last Year's Incoming Scores	Current Outgoing Scores
HET	0.7854***	0.7325***	0.8145***	0.5601***
	(0.0208)	(0.0275)	(0.0216)	(0.0479)
HET2	-0.0105***	-0.0080**	-0.0087**	-0.0060
	(0.0038)	(0.0038)	(0.0038)	(0.0039)
HET3	-0.0226***	-0.0213***	-0.0245***	-0.0147***
	(0.0024)	(0.0025)	(0.0024)	(0.0029)
HET4	-0.0001	-0.0004	-0.0006	-0.0001
	(0.0007)	(0.0007)	(0.0007)	(0.0007)
predicted school-level HET	0.0703***	0.1309***	0.0370	0.3281***
	(0.0232)	(0.0309)	(0.0240)	(0.0546)
big fish	0.0312***	0.0440***	0.0170*	0.1009***
	(0.0086)	(0.0106)	(0.0090)	(0.0167)
small fish	-0.0741***	-0.1051***	-0.0563***	-0.2078***
	(0.0169)	(0.0197)	(0.0170)	(0.0304)
art students	-0.1119***	-0.1212***	-0.1198***	-0.1251***
	(0.0056)	(0.0056)	(0.0056)	(0.0057)
female	0.0523***	0.0515***	0.0520***	0.0509***
	(0.0078)	(0.0078)	(0.0078)	(0.0078)

Table II.7 Continued

	Dependent variable: College Entrance Test (CET) score			
	(1)	(2)	(3)	(4)
VARIABLES	Current Incoming Scores	Average (Past 3 Years)	Last Year's Incoming Scores	Current Outgoing Scores
private school	0.0139	0.0207	0.0156	0.0297*
	(0.0155)	(0.0157)	(0.0157)	(0.0159)
paid high price	0.0253**	0.0233**	0.0212*	0.0271**
	(0.0110)	(0.0110)	(0.0110)	(0.0110)
rural*female	0.0949***	0.0832***	0.0790***	0.0920***
	(0.0150)	(0.0153)	(0.0153)	(0.0155)
suburb*female	0.1143***	0.1057***	0.1039***	0.1079***
	(0.0107)	(0.0108)	(0.0108)	(0.0108)
suburb	-0.0174	-0.0020	-0.0186	0.0329**
	(0.0117)	(0.0127)	(0.0122)	(0.0149)
rural	-0.1658***	-0.1575***	-0.1690***	-0.1322***
	(0.0085)	(0.0089)	(0.0086)	(0.0105)
Observations	47,267	45,367	45,367	45,171
R-squared	0.6934	0.6890	0.6889	0.6878

Notes: Standard errors in parentheses; ***p<0.01, **p<0.05, *p<0.1. The column titles are the scores used as the predictor of school quality.

II.5 Conclusion

This paper considers whether or not students benefit from being a big fish in a small pond, or from dropping down from the best possible school they could attend in order to gain a position higher in the relative score distribution of their school. Using student-level test score data from a large metropolitan city in China, we find evidence that while both individual student quality and school quality matter, there is a separate "big fish" effect. Students who a classified as big fish in a small pond see an increase in their test scores of about 0.6 standard deviations on average. Further, this effect increases as the distance increases between the cutoff score of the student's best possible school and the cutoff of the student's actual school (or as the student drops down from his best possible school). This suggests that in considering the school choice decision, students should consider not only school quality, but also their positions relative to their potential peers.

CHAPTER III

PEER QUALITY AND THE ACADEMIC BENEFITS TO ATTENDING BETTER SCHOOLS

III.1 Introduction

A common feature of educational systems around the world is that students sort into high schools and colleges on the basis of ability. In the United States, the sorting at the high school level takes place largely by families moving across neighborhoods and school attendance zones, while in college and in the much of the world the sorting is based on demonstrated academic performance. Across all of these contexts, students and their families exhibit strong revealed preferences for attending more selective high schools and colleges composed of higher achieving peers.

However, while recent research has documented significant returns to college quality (e.g., Andrews, Li, and Lovenheim, 2012; Canaan and Mouganie, 2015; Hoekstra, 2009; Saavedra 2009; Zimmerman, 2014), the literature on the returns to high school quality is less con- clusive. Many recent studies find that attending middle and high schools with significantly higher-performing peers does not improve academic performance (Abdulkadiroglu, Angrist, and Pathak, 2014; Clark 2010; Dobbie and Fryer, 2014; Lucas and Mbiti, 2014; Zhang, 2014). In contrast, others find that attending more selective schools does result in benefits (Berkowitz and Hoekstra; 2011; Clark and Del Bono, 2014; Ding and Lehrer, 2007; Jackson, 2010; Park, Shi, Hsieh, and An, 2015;

Pop-Eleches and Urquiola, 2013). In addition, in some cases the benefits are relatively modest. For example, Pop-Eleches and Urquiola (2013) find that students in Romania who attend schools with peers that are one standard deviation better score only 0.1 to 0.2 standard deviations higher on the national high school exit exam. Overall, the lack of consistent evidence of meaningful effects presents a puzzle. The purpose of this paper is to address this puzzle by examining the returns to high school quality across a range of high schools with different levels of selectivity, and in a context in which we can use measures of other school inputs such as class size and teacher quality to speak to potential mechanisms.

To do so, we apply a regression discontinuity design that exploits a unique institutional feature of the educational system in China. All students in China who wish to attend high school must sit for a national entrance exam, performance on which determines high school eligibility. While some students score barely above these cutoffs, others score just below. Intuitively, we compare the college entrance exam performance of these students to each other, which enables us to distinguish the effect of attending more selective high schools from unobserved confounding factors such as ability and motivation.

This approach has two primary advantages. The first is that the institutional context and administrative student-level data are ideally suited for providing credible estimates of the returns to high school quality. This is in large part because the high school admission thresholds are determined only after students take the exam, making it difficult—if not impossible—for students to manipulate where they are relative to the cutoff. As a result, it is difficult to imagine a scenario in which students barely above and below the threshold are not otherwise similar to each other. In addition, the institutional framework we study provides for significant variation in high school quality; most students barely above the cutoff choose not attend the more selective school, and very few students just below the cutoff are able to attend that school. This contrasts with other contexts where student preferences and admission rules create much smaller discontinuities in attendance, which limits the interpretation of the estimates to a relatively small set of compliers. Finally, because of the high demand for attending four-year colleges, and because the college entrance exam is the determining factor for university admissions, students are highly incentivized to take the exam. As a result, there is little scope for selection into test-taking to bias the estimates. More importantly, because both teachers and students face strong incentives to do well on the college entrance exam, performance on it should be a good measure of whether students learn more at more selective schools. This contrasts with other settings, where observed performance outcomes may not be good measures of what students and teachers hope to achieve during high school.

The second main advantage of our study is that we are able to estimate gains in academic performance across a range of high schools that differ in quality but are within the same educational context. Thus, we are able to estimate returns not only for elite schools, but also for less selective schools in the same region, all of whom have discontinuously higher- achieving peers than the schools just below the cutoff. More importantly, this setting enables us to exploit differences across these cutoffs with respect to returns and education inputs in order to provide evidence on the importance of potential mechanisms. In particular, our data include measures of both class size and teacher quality as well as peer quality. Our measure of teacher quality is the concentration of "superior" teachers. This is the top rank of teachers in China, and the only one that cannot be earned based on credentials such as advanced degrees. Instead, it is based on rigorous evaluations of performance, a significant component of which is student performance on the college entrance exam. Data on these potential mechanisms turn out to be important, as the results we document are difficult to reconcile with the hypothesis that peer quality is responsible for returns to school quality. In addition, the fact that we are able to exploit the heterogeneity in returns and inputs within the same educational context reduces worries that any differences in returns are due to differences in behavioral responses, such as those documented by Pop-Eleches and Urquiola (2013).[1]

Results indicate that across the full range of high schools, there are few academic benefits to attending more selective high schools. Specifically, using a stacked RDD approach similar to that of Pop-Eleches and Urquiola (2013), we document that being

---

[1] The focus on mechanisms underlying the returns to high school quality is one difference between our paper and that of Park, Shi, Hsieh, and An (2015), who estimate the returns to attending magnet high schools in a different province of China. In addition, there are other important differences. Our data are from the Ministry of Education, while theirs are from schools contacted directly who agreed to share data. Perhaps as a result, we observe college entrance exam scores for 91 percent of students, which is consistent with the officially reported range using aggregate data, compared to 62 percent for Park, Shi, Hsieh, and An (2015). We also observe significantly higher compliance in our data for students barely above and below the threshold. In our sample, threshold-crossing at the Tier I cutoff is associated with a roughly 70 percentage point increase in the likelihood of attendance, which as twice as large as the increase in Park, Shi, Hsieh, and An (2015).

barely admitted to a more selective school is associated with an average of a one-fifth standard deviation increase in peer quality. Similarly, we show that there are meaningful increases in peer quality across different admission thresholds throughout the range of high schools. However, we find no evidence that attending these schools with higher-ability peers leads to improved college entrance exam performance, on average. In contrast, when we focus on the return to attending elite Tier I schools, as other recent studies have done in the US (Abdulkadiroglu, Angrist, and Pathak, 2014; Dobbie and Fryer, 2014), we find significant returns. Specifically, we document that attending Tier I schools leads to a 0.16 standard deviation increase in exam performance. Given this exam is the primary determinant of admissions to universities in China, these gains lead to significant increases in students' ability to attend four-year colleges, which has been shown to have substantial returns in China (Giles, Park, and Wang; 2015).

Interpreted in the context of peer quality, these findings present a puzzle. That is, while we document that threshold-crossing is associated with significant increases in peer quality across all schools even outside of the Tier I threshold, the only returns come from attending Tier I rather than Tier II schools. We further document that these findings are difficult to reconcile even by the presence of non-linear peer effects, or by the possibility that only students at top schools are incentivized to do well on the exam. We do so by documenting that while there is a significant discontinuity in peer quality across admission thresholds within Tier I schools, there is no evidence of improved performance.

Instead, we find that these results are most consistent with the hypothesis that returns to high school quality are caused by teacher quality, rather than peer quality. Specifically, we find that the only discontinuity in access to prestigious superior teachers is at the Tier I/Tier II threshold. In contrast, while there are large discontinuities in peer quality across all other thresholds—including within Tier I, within Tier II, and across Tier II and III—barely admitted students in those schools do not have more access to superior teachers. Similarly, we find no evidence that the pattern of results could be due to differences in class size; if anything, class size is larger for students who attend Tier I versus Tier II schools.

The finding that teacher quality, rather than peer quality, is likely responsible for returns to attending more selective schools is consistent with previous estimates on the value-added of superior teachers in China. Hannum and Park (2001) estimate that superior teachers improve math and language scores by 0.14 and 0.44 standard deviations, respectively, relative to the lowest ranked teachers. By comparison, a back-of-the-envelope calculation suggests that if superior teachers in our context had a value-added that was 0.13 to 0.17 standard deviations better than average, the increased access to those teachers at Tier I schools would explain all of the positive return we estimate. We view this as plausible, and consistent with recent evidence in the US highlighting the importance of teacher quality (Chetty, Friedman, and Rockoff, 2014), as well as with Jackson (2013), who finds that peer achievement can only explain a small part of the returns to selectivity in Trinidad and Tobago.

The results of this study may also help explain the mixed findings in the literature, all of which has documented significant increases in peer quality, but only some of which reports evidence of performance gains. The finding here of substantial differences in the returns to school quality within the same educational context suggests that there should be an increased focus on understanding and measuring why school quality matters. This is important because different mechanisms have substantially different policy implications. For example, if gains due to selective schooling were due to peer effects, there would be limited scope for enabling more students to benefit from school quality. On the other hand, if gains are driven by differences in teacher quality, then it may be possible to extend the benefits of attending better schools to more students, without reducing returns to others. Results in this study are more consistent with this latter interpretation, since the only positive returns to high school quality occur when there is also a significant increase in teacher quality.

III.2 The Chinese Education System

III.2.1 Overview of Schooling in China

Children in China generally start elementary school at the age of six or seven. After spending six years in elementary school, children then move on to the first part of middle school, which lasts three years (7th to 9th grade) and completes the nine year national compulsory education requirement. Graduates from junior middle school then choose to pursue either vocational or traditional schooling. The traditional education path involves participating in the second part of middle school, which is equivalent to US high schools. Three years of high school are then followed by higher education (university/college) for those who are willing and able to do so. Only high school graduates are eligible for university; vocational graduates are not.

In China, elementary and middle school education are both free and compulsory. On the other hand, high school education is neither compulsory nor free. However, in most parts of China, the majority of high schools are public and charge relatively low tuition. For example, in the province we study, public high school costs around $200 per year, and can be less if family income is below certain amounts. Around 60 percent of junior middle school graduates in our province attend high school, while the rest attend vocational schools. Less than 5 percent of students attend private high schools, which are generally not as good as public schools.

There is vigorous competition amongst middle school students to enroll at the selective high schools, and admissions are most competitive at the highest-ranked high schools. Ad- mission to high schools is based on a city-level entrance exam called the Zhongkao, or the HET, which is comprised of seven subjects. These subjects are Chinese language, Mathematics, English language, Physics, Chemistry, Political Science and Physical Education. The weighted sum of these seven subjects is the one and only criterion for high school ad- mission for most students; the only (rare) exceptions are students with special talents, such as athletes. The HET is graded out of a possible 790 points.[2]

During high school, students usually choose an academic concentration (Arts or Science) at the beginning of their junior year (2nd year) in high school. Some college majors only admit students from one path and others accept both, so this choice can be a combination of personal preference and comparative advantage.

Similar to the high school admission process, university admission decisions are made almost entirely on the basis of performance on the college entrance exam, called the Gaokao or the CET. This exam is taken after three years of high school, and is required of all students who wish to attend college. High school students concentrating in arts take an exam that includes Chinese language, Mathematics for arts students, English language and a comprehensive arts test consisting of Political Science, History, and Geography. Students concentrating in sciences take an exam that includes Chinese language, Mathematics for science students, English language and a comprehensive science test consisting of Physics, Chemistry, and Biology. The exam for both tracks is graded out of a possible 750 points.[3]

In contrast to high school, students in China are free to attend any university—regardless of location—conditional on meeting that university's threshold score. In

---

[2] Chinese, Math and English are each graded out of a possible 150 points, while Politics, Physics and Chemistry are each graded out of 100 points. Physical Education is graded out of a possible 40 points.
[3] For the science track, Chinese, Mathematics for the sciences, and English are each graded out of 150 points, while the science comprehensive test is graded out of 300 points. For the arts track, Chinese, Mathematics for the arts, and English are each graded out of 150 points, while the arts comprehensive test is graded out of 300 points.

addition, eligibility to attend any four-year college in China is determined by specific thresholds set by each province. As a result, students are heavily incentivized to perform well on the CET. Indeed, the desire to do well on this exam is the main reason for the competitive admissions process into high schools, as students hope to position themselves to do well on the college entrance exam and thus attend a selective university.

III.2.2 High School Choice Mechanism

High school admissions for the province we study is centrally operated by each city's education administrators. In early June, students fill out application forms indicating their ordered preference of high schools. These students then take the high school entrance exam in mid June. High schools predetermine how many students they wish to admit for that year and grant admission based on students' preferences and test scores. Most provinces, including ours, use an admission procedure similar to the Boston Mechanism. In the first round of admissions, each high school only considers students who list them as their first choice. Students with HET scores above a certain threshold are accepted and the rest are rejected and placed in a pool of candidates to be considered by the next high school on a student's list. Only in the event that a high school still has any remaining slots after the first round will it consider admitting students who list them as their second or third choice. Once a student is granted admission by any high school, the selection process ends for that student and he/she is not to be considered by any other high school.

For illustration, suppose school A plans to enroll 100 students for that academic year. Further, suppose that there are 80 students who indicate their first preference is to join that school. School A will first admit those 80 students, then proceed to rank students who listed A as their second choice—conditional on not yet being admitted by their first choice. If there are more than 20 of those applicants, school A will take the 20 highest scoring students. If admission slots remain, then School A proceeds to fill the rest of their seats with students who list A as their third choice, and so forth. In the more likely scenario that there are more than 100 students who list School A as their first preference, officials select the highest scoring 100 students. The lowest admitted student's score—regardless of preference order— is the official cutoff score for school A for that year. High schools go through this process simultaneously, as each student can be admitted by at most one school.[4]

[4] Public high schools in our sample are allowed to designate around 10% of their seats as "high priced". Students enrolled through the high-priced channel pay a one-time fee to the school upon registration, though they receive the same education as the other regular students. This one-time fee is set by the schools and revealed to students before they apply. In urban areas it is usually around 40,000 Yuan (6,600 USD) and 20,000 Yuan (3,300 USD). Schools allocate these high-price slots in a separate but otherwise similar process as that used to allocate the other slots. For example, suppose school A plans to set 90 regular seats and 10 high-priced seats. Then A (regular) and A (high-priced) independently go through the high school admissions process as described above. Students decide which schools, regular or high-priced, to apply at the same time, before the test. Students can even apply to both regular and high-priced of the same school. Thus, all schools with high-priced seats, will have two cutoffs—one for regular students and one for high priced students—and this information is released to the public by both education officials and the media. In our analysis, we keep in the sample all individuals who entered high school through this "high priced" process, though those students are likely "non-compliers" and thus contribute little to the variation we use to identify effects.

To ensure smoother and more transparent school-student matching, schools are divided into four groups by the city education department. These groups are defined in advance of student applications, and the groupings are made public to all students. We call these groups “tiers“ as they divide the schools with respect to quality/selectivity. The best schools are called Tier I schools, the second-best are Tier II, and so on. The composition of each tier is quite stable over time, though sometimes schools change tiers from one year to another to reflect changes in quality. In addition, there are also differences in school selectivity within tiers as well as across tiers.

All schools are also ranked nationally according to the following designations (from best to worst): National demonstrative high schools (“Guojia Shifanxing Gaozhong”), Provin- cial first class schools (“Shengyiji Xuexiao”), Municipal first class schools (“Shiyiji Xuexiao”), District level first class schools (“Quyiji Xuexiao”) and ordinary schools (“Putong Zhongxue”). All Tier I schools in our sample are designated as national demonstrative high schools. This ranking system was introduced by the Ministry of Education during the State’s ninth “Five-Year Plan” period (1996-2000). To earn this title of “National Demonstrative High School, a school must meet certain criteria regarding curriculum design, school facilities, teacher quality and student performance.[5]

[5] For example, according to the national demonstrative school assessment scheme on the provincial department of education website, at least 30% of the teachers must have either a graduate degree or superior teacher title; student crime rate must be lower than 1%; and at least 25% of students must meet the provincial elite college cutoff in the CET and 60% must meet the four-year college cutoff.

Schools in the first tier begin the admissions process. After Tier I schools fill all their seats, Tier II schools will start admitting students, then Tiers III and IV. Accordingly, students list their preferences by tier. For each tier, a student has four ordered school choices. Importantly, because there are fewer than four Tier I schools in the districts we analyze in this study, students are able to list and rank each of the Tier I schools. The order of choice is important as most competitive schools fill their slots solely with students who have them listed as their first choice. Students generally understand this and as a result most list their preferences by perception of school quality.

School choice is different for students from different parts of the city we analyze. Specifically, the city is divided geographically into twelve administrative districts, which define the region in which students have choice regarding high school. Of the twelve districts, eight of them are mostly urban areas and are relatively small. Students from these eight districts can go to high schools in their own district but also have access to schools in the other seven urban districts. Specifically, a student residing in an urban area can choose from almost all urban Tier I schools—regardless of district—in addition to schools in Tier II to IV from their own district. On the other hand, students from suburban districts on average have a choice set of only two Tier I schools and 11 Tier II through IV schools. Further, students residing in the four suburban districts can only choose high schools in their own district. As a result of the more limited choice sets facing students in the four suburban districts, the admission system generates much more significant discontinuities with respect to school selectivity and ability levels. The students in these districts also have much more uniform preferences over school quality, given the significant differences across the limited set of schools. For these reasons, we restrict our analysis to these suburban districts.

III.2.3 Type of Teachers in High School

A unique and important aspect of high school education in China is the clear distinction of teachers by rank. There are different titles (ranks) for high school teachers, and salaries increase with these ranks. The three professional ranks for all public school teachers, regardless of class level are “elementary”, “intermediate” and “superior”. Further, within the intermediate rank, there are two smaller categories; “second class” and “first class”.

One automatically becomes an “elementary” rank teacher upon employment in the teaching sector. However, if that person holds a master’s degree, then they start at the intermediate second class rank. Teachers with a doctoral degree start at the intermediate first class level. After two years as an elementary rank teacher, a teacher then applies for promotion to the intermediate second class level. After four years within this rank, they are able to apply for the intermediate first class rank. Finally, after achieving a first class rank and after a period of no less than five years,[6] a teacher is permitted to apply for the superior teacher rank.

---

[6] Master degree holders can apply for promotion to first class rank after two years instead of the usual four. Doctorate degree holders only need to have two years of teaching experience to be eligible for superior teacher promotion.

After a teacher applies for promotion, and after the approval of the school they work for, city education officials put together a committee to start an evaluation process assessing a teacher's performance along several dimensions such as teaching, publication level, character, and various other aspects in his/her field. A committee will quantitatively grade a candidate on these aspects.[7] For example, having a PhD degree is worth 3 points, a Masters degree is worth 2 points, and a Bachelor's degree is worth 1 point. There are five categories and 100 total points: Degree (3 points), Tenure (7 points), Experience in Current Position (22 points), Performance in Current Position (38 points), Research Papers (10 points) and Awards and Contribution (20 points).[8] Within the Performance category, there are four sub-categories, one of which is based on teaching outcomes. In this subcategory, a candidate can score up to 13 points if their students' average test scores are high (8 points), their students improve a lot on a particular subject (3 points) and they forms a special and effective teaching style (2 points). The assessment ends with an oral exam. If more than 2/3 of the committee members vote for approval, then a teacher will be approved for the promotion. Similar to tenure in the U.S., once a teacher is promoted to a higher rank, they generally do not get demoted. For example, in our sample during the time period we study, no teachers were demoted for failing to meet certain assessment tests.

Salaries differ by rank, so teachers have an incentive to get promoted. Teacher salaries in China consist of two main parts: 1) state (base salary) and 2) local (city,

[7] http://www.gzedupg.com/zhicheng detail.php?pid=10&id=398.
[8] The assessment table can be found at: http://www.gzedupg.com/download/sjh2008362fb2.xls

district and school level). The local part of teacher salary varies extensively from city to city and even from school to school and can be based on performance. The state base salary is determined by one's professional rank and title ("Gangwei Gonzi") as well as years of service ("Xinji Gonzi") and has a nationwide set of standards. For instance, superior teachers receive an additional "Gangwei Gonzi" salary of 1180 RMB ($190) to 930 RMB ($150) per month. On the other hand, first class teachers receive an additional "Gangwei Gonzi" salary of 780 RMB ($126) to 680 RMB ($110) per month. Superior teachers' "Xinji Gonzi" portion of their salary starts from level 16 (317 RMB per month), while first class teachers start from level 9 (181 RMB per month).

An important question is whether teachers of higher rank in China have higher value- added when it comes to the college entrance test scores of their students. While our under- standing of the promotional process leads us to believe that this would be true in particular for superior-rank teachers, to our knowledge there are two empirical studies that speak directly to this question. Using lottery data from Beijing middle schools, Lai, Sadoulet, and de Janvry (2011) show that teacher rank is highly correlated with estimated school fixed effects, suggesting that a significant part of school quality is due to teacher rank. In addition, Han- num and Park (2001) find that teachers with superior rank increase math and language test scores by 0.14 and 0.44 standard deviations, respectively. They conclude that the teacher quality ranks reveal significant information about teacher quality that is not contained by measures such as the teacher's degree attainment and years of experience.

III.3 Data

We use student-level administrative data from a large capital city of a densely populated province of more than 7,000 square kilometers in China. As a condition of using the data, we are unable to reveal the name of the province and city. The city has a population of more than 10 million and a per capita GDP of more than $20,000. The two suburban areas we are looking at in this paper have a total population of more than 2 million. GDP per capita is around $16,000, which is lower than the urban part of the city but still higher than the national average.

Our data come from the education bureau authorities of the city. The authorities merged student data of those who took the High School Entrance Test (HET) and attended one of the traditional high schools in 2007 with those who took the College Entrance Test (CET) in 2010, resulting in a sample size of 49,674 students.[9] For each student, we observe both their HET and CET scores and some student characteristics including the middle school and high school attended, gender, age, middle and high school district, and parents' occupations. Because some high school students do not take the CET, in Section 5.3 we test for selection into taking this exam and perform bounding exercises to ensure our estimates are not biased by selection into test-taking.

Our data only contain individuals attending traditional high school since those attending vocational schooling are not generally eligible to take the college entrance examination. Further, we restrict our sample to suburban districts, where students have a

[9] According to official records, a total of 59,591 students registered for the CET in this city in 2010, which includes the current high school seniors as well as those who already graduated and wished to take the test again.

limited choice set compared to students in urban districts. In addition, our analysis focuses on the two suburban districts that have at least one school that is exclusively Tier I.[10] Our final sample consists of 15,367 students taking the high school entrance exam (HET) in the year 2007 and the college entrance exam (CET) in 2010.

Data on school and teacher characteristics were collected from government reports and official school websites. These data include the size of the schools, which include the geo- graphic size of the school, the number of students, classes, teachers and superior teachers. This information was collected from government reports as well as the official school web- sites' "school overview section for most schools and recruitment advertisement pamphlets for the few schools that do not have a website. We link these data to our student data using school identifiers.

Within the two suburban areas we analyze, students generally have at most 15 high schools to choose from. The tier of each school is clear and widely known. The main determinant of which high school is attended is the score on the high school entrance exam. We observe detailed administrative data on test scores by subject and the eventual high school attended by the student. As a result, we are able to measure school peer quality by calculating the average score on the HET for students in each school.

---

[10] There are four suburban areas where students must attend high schools within the same district. One of them does not have any Tier I schools, while another has five high schools that are simultaneously classified as Tier I and Tier II. Because much of our analysis is focused on the returns to Tier I schools, we leave both of these districts out of the sample, though later on we perform robustness checks showing that our estimates are largely unchanged when we include the latter district, regardless of how we classify those five Tier I/Tier II schools.

Table III.1 Descriptive Statistics

	(1) Whole Sample	(2) Tier I Schools	(3) Tier II Schools	(4) Tier III Schools	(5) Tier IV Schools
High school entrance exam scores	617.32	674.07	607.5	545.9	495.21
	(58.40)	(27.98)	(33.99)	(38.61)	(31.92)
College entrance exam scores	490.98	574.46	471.74	395.22	342.47
	(98.28)	(55.8)	(79.03)	(77.81)	(66.88)
Proportion female	0.54	0.55	0.53	0.54	0.50
Proportion majoring in arts in high school	0.48	0.33	0.50	0.65	0.68
Proportion private schools	0.010	0.006	0.015	0.018	0
Eligible for four year college	0.43	0.85	0.28	0.056	0.016
Eligible for elite college	0.088	0.243	0.012	0.002	0
Proportion female teachers	0.55	0.55	0.56	0.57	0.52
School Size	127.27	145.66	122.45	102.72	83
	(42.26)	(47.71)	(34.70)	(27.90)	(0)
Number of students per teacher	15.65	16.20	15.43	15.75	11.8
	(2.86)	(1.34)	(3.24)	(3.84)	(0)
Number of students per superior teacher	216.25	45.24	313.81	404.59	66.86
	(311.48)	(8.99)	(388.83)	(185.51)	(0)
Number of schools	25	4	12	8	1
Number of Students	11125	3744	5367	1696	309

Notes: Data taken from two rural districts in the Province for students taking the high school entrance exam in 2007. Standard errors (for non-binary variables) in parentheses.

The main outcome of interest is a student's total score on the college entrance exam. In addition, we also examine eligibility to attend a four year college. We can do so because eligibility for entry into a four year college is centrally determined by whether a student crosses the lowest threshold score imposed by a specific university. This threshold is common to all students in a province regardless of which city they reside in within the province. As a result, while we do not observe the university a student ultimately attends, we are able to determine whether students are eligible to enter any four year college using their final CET scores.

Descriptive statistics for all students who sat for the 2007 high school entrance exam are reported in Table III.1. These statistics are reported for the whole sample and by high school tier. The average scores on the HET and CET are 612 and 492 points, with standard deviations of 61 and 97, respectively. These scores increase with the level of tier as one would expect. Just over half of the high school students (53%) are female. For the full sample, 48% of the students choose to major in arts, though that figure ranges from 33% in Tier I to 68% in Tier IV. Very few students attend private high schools (1%). A total of 57% of students attend high schools that include only the high school grades, while the others contain both middle and high schools. 44% of students in our sample are eligible to go to a four-year college. This number is as high as 67% for students attending Tier I high schools and drops to 28%, 7% and 1% for Tiers II to IV. Higher tier schools tend to be larger in size. The ratio of students to teachers is relatively stable across all tiers at around 15 to 16 students per teacher, except for the one Tier IV school which has a ratio of 12. More selective schools have a significantly lower student to superior teacher ratio; Tier I schools have, on average, 45 students per superior teacher, compared to 313 for Tier II schools. Finally, 53% of all teachers in our sample are female, which is roughly constant across tiers.

III.4 Identification Strategy

III.4.1 Single Cutoff: The Academic Return to Attending Tier I Schools

As mentioned earlier, high schools in the province we are analyzing are divided into four tiers with the first tier containing the best set of high schools within a district or town. Accordingly, we use a regression discontinuity design (Lee and Lemieux, 2010; Imbens and Lemieux, 2008) to estimate the causal impact of elite high school attendance (defined as going to a Tier I high school) on college entrance exam scores and college attendance. The key identifying assumption underlying an RD design is that all determinants of future outcomes vary smoothly across the Tier I high school admissions threshold. This is likely to hold, as precisely manipulating the overall exam score would be extremely difficult, if not impossible. This is because the cutoff scores for each high school are only determined after the exams are administered and graded. These cutoffs are determined based on high school applicants' percentile ranks, which are only calculated after the tests are graded. As a result, students and graders do not know where the admission thresholds for each school lie until after the test is taken and graded. In addition, graders do not observe any identifying information on students, so it is not possible for them to artificially increase the grades of certain students.

All students in our data attend a school in one of two suburban towns.[1] Accordingly, we have two Tier I cutoffs in our data—one for each town. In order to summarize the effects of attending an elite school, we run regressions that pool data across both towns. Formally, consider the following reduced-form regression:

$$Y_i = \alpha + h(S_i) + \tau D_i + \delta X_i + \varepsilon_i, \qquad \text{(III.1)}$$

where the dependent variable Y is the outcome of interest. D is a dummy variable indicating whether a student i crosses the town-specific score threshold for attending a Tier I high school.[2] S represent student high school entrance test (HET) scores in 2007 measured in points relative to the cutoff score of each town, relying on the fact that S represents the distance between each town's cutoff and the HET score of each student in that town. Formally, $S_i = grade_i - \overline{grade_z}$ for all individuals within a town facing a common Tier I cutoff z. The function h(.) captures the underlying relationship between the running variable and the dependent variable. We also allow the slopes of the fitted lines to differ on either side of the admissions threshold by interacting h(.) with the treatment dummy D. X is a vector of controls that should improve precision by reducing residual variation in the outcome variable, but should not significantly change the treatment estimate if our identifying assumption holds. The term ε represents unobservable factors affecting outcomes. Finally, the parameter τ gives us the average effect of having the opportunity to access a Tier I high school for each outcome of interest.

---

[1] As mentioned earlier, students may only attend high school in the suburban town in which they reside.

[2] We have two thresholds, with each representing a different town.

Table III.2 “Stacked RD” Estimates for Attending Better Schools Across All Admission Cutoffs

Bandwidth	2.5 CCT (1)	2 CCT (2)	1.5 CCT (3)	1.25 CCT (4)	CCT (5)	0.75 CCT (6)
Panel A:						
Discontinuity in high school peer quality	.189***	.204***	.149***	.188***	.187***	.205***
	(.025)	(.028)	(.021)	(.024)	(.027)	(.031)
With Controls	.187***	.203***	.148***	.188***	.187***	.204***
	(.023)	(.026)	(.020)	(.023)	(.026)	(.030)
Observations	33,414	26,806	20,221	17,369	13,570	9,619
Panel B:						
Discontinuity in CET exam scores	−.002	−.005	−.017	−.009	−.007	.006
	(.016)	(.020)	(.013)	(.016)	(.020)	(.025)
With Controls	−.006	−.009	−.019	−.014	−.014	−.004
	(.016)	(.019)	(.013)	(.016)	(.019)	(.024)
Observations	66,530	53,930	41,599	34,334	27,777	21,135
Panel C:						
Discontinuity in likelihood of enrolling in 4-year college	−.001	−.004	−.003	−.001	−.001	.009
	(.008)	(.005)	(.006)	(.007)	(.009)	(.012)
With Controls	−.002	−.004	−.004	−.001	−.003	.005
	(.007)	(.005)	(.006)	(.007)	(.009)	(.011)
Observations	70,493	57,330	42,524	37,029	29,682	22,065
Score Polynomial	Two	Two	One	One	One	One

Notes: Sample includes students who took the college entrance exam in the year 2007. Controls include: Age, gender, district fixed effects and middle school fixed effects. Optimal Bandwidth selected using the CCT bandwidth selector proposed in Calonico et al. (2015). Optimal BW = 14 for high school peer quality regressions. Optimal BW = 29 for CET exam regressions. Optimal BW = 31 for likelihood of enrolling in four year degree regressions. Since we observe individuals with multiple cutoffs, we cluster at the student ID level. *** p <0.01 ** p <0.05 * p <0.1

In our analysis, we specify h(.) to be a linear function of S and estimate the equation over a narrower range of data, using local linear regressions with a uniform kernel. This approach can be viewed as generating estimates that are more local to the threshold and does not impose any strong functional assumptions on the data. As a result, the preferred specifications in this paper are drawn from local linear regressions with the optimal bandwidths chosen by a robust data driven procedure as outlined in Calonico, Cattaneo and Titiunik (2014)—henceforth CCT. We also present results for a variety of bandwidths relative to the optimal bandwidth as has become standard in the RD literature (Lee and Lemieux, 2010). Further, standard errors are clustered at the high school score (HET) level, as suggested in Lee and Card (2008).

III.4.2 Multiple Cutoffs: The Academic Return to Attending Better Schools

Within each of the two suburban towns in our sample, we rank schools according to their posted admissions cutoff score for that year (2007). This yields 23 quasi-experiments as each cutoff results in a potential RD analysis.[1] Following Pop-Eleches and Urquiola (2013), we focus on regressions that pool data across all school entry cutoffs. Specifically, we stack the data such that each student within a certain town serves as a separate observation for each cutoff.[2] Due to repeated observations, we

[1] We have an average of around 12 schools for each town resulting in 11 different cutoffs within each town.

[2] For instance, our smallest town has 12 different schools, leading to 11 separate cutoffs. This town also contains 4,025 students. For that town, our procedure produces a dataset of (4,025 × 12) 48,300 observations.

cluster our standard errors at the individual level. Formally, our reduced form regression from this procedure takes the following form:

$$Y_{iz} = \alpha + h(S_{iz}) + \omega D_{iz} + \varphi X_i + \varepsilon_i. \tag{III.2}$$

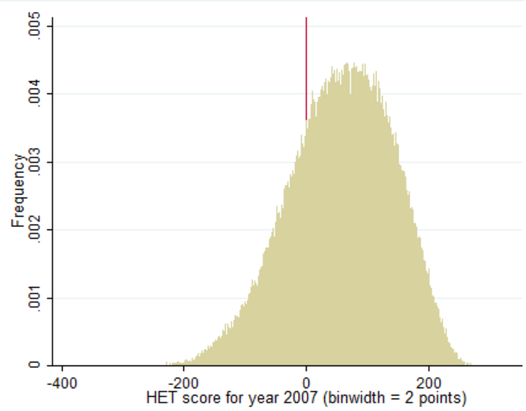
Here the subscript i still refers to students and the subscript z refers to all possible high school cutoffs facing an individual within a town (i.e. z= 1,...,H-1; where H represents the total number of high schools in that town ordered from worst to best based on their respective cutoff scores). ω gives us the ITT estimate of having the opportunity to go to a better school, regardless of tier. Further, the running variable is defined as $S_i = grade_i - \overline{grade_z}$ for all individuals within a town facing numerous cutoffs z. As a result, equation (III.2) takes the same form as equation (III.1) except for the fact that each individual can be observed multiple times depending on his/her relative position to a high school cutoff. However, regressions restricted to students scoring close to the cutoffs rarely use student-level observations more than once.

III.4.3 Tests of Identification

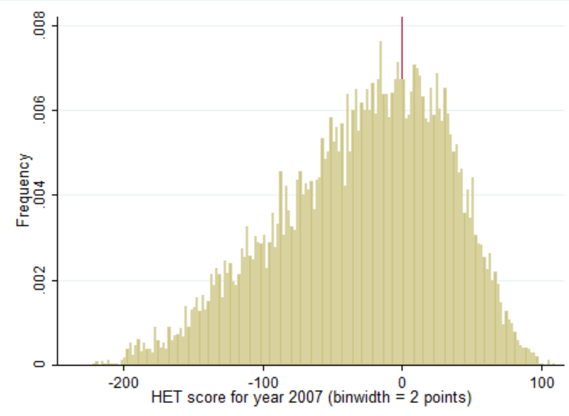
As described above, given the nature of the school assignment mechanism and the way in which it is implemented, we find it unlikely that students are able to manipulate the assignment variable in such a way that would invalidate the research design. However, we still provide empirical tests in order to assess whether the data appear consistent with the identifying assumption that no other determinants of achievement vary discontinuously across the threshold.

Figure III.1 Testing the Validity of the RD Design for Both Empirical Strategies

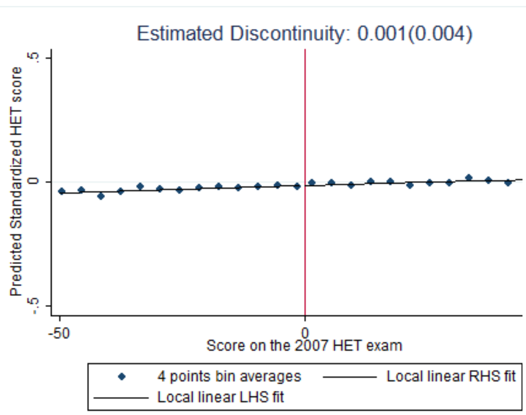
Attending Better Schools (All Cutoffs) | Attending a Tier I school



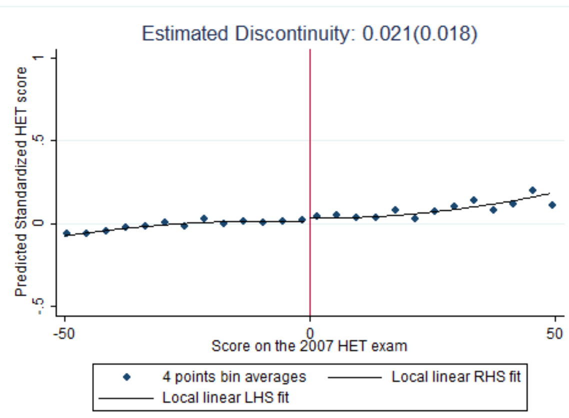
(a) Distribution of HET scores



(b) Distribution of HET scores



(c) Smoothness of baseline covariates



(d) Smoothness of baseline covariates

Notes: Sample includes students who took the HET exam in the year 2007.Bins for histogram represent an average count of 2 score points.Predicted score based on the following controls: sex, gender, district fixed effects, middle school fixed effect.

First, we ask whether there is any evidence of bunching around the admission threshold. Under our identifying assumption, there should be no such bunching. In contrast, if students or graders could manipulate scores relative to the cutoff, we might expect to see too few students just short of the cutoff, and too many students barely exceeding the cutoff.

Results are shown in Panels A and B of Figure III.1, which show the density function for the stacked RDD across all admission cutoffs as well as the that for only the Tier I admission threshold. Both show no evidence of bunching around the cutoff, consistent with the identifying assumption.

In addition, we also test whether observed determinants of achievement are smooth across the threshold. If the identifying assumption holds, we expect all such variables to vary smoothly across the admission thresholds. On the other hand, if students or graders are able to manipulate scores around the cutoff, then we might expect to see evidence of different types of students on either side of the cutoff. Covariates in our data set include age, gender, and district and middle school fixed effects.

Rather than focusing on these covariates individually, we instead use those covariates to predict college entrance test scores for each student. We then ask whether those predicted scores are smooth across the cutoff.[3] We do this in part because using this weighted average of characteristics corresponds most closely to what we care about - whether underlying ability to do well on the college entrance exam varies smoothly across the cutoff. In addition, the predicted performance measure can more easily quantify the role of middle schools and district attended by the students. Results are shown in Panel C of Figure III.1, and indicate that there is little evidence that underlying student ability varies discontinuously across the threshold. Estimates shown in Appendix

[3] In Table A1, we also show estimates for age and gender separately.

Table A1 are also close to zero and statistically insignificant across a range of bandwidths.

III.5 Results

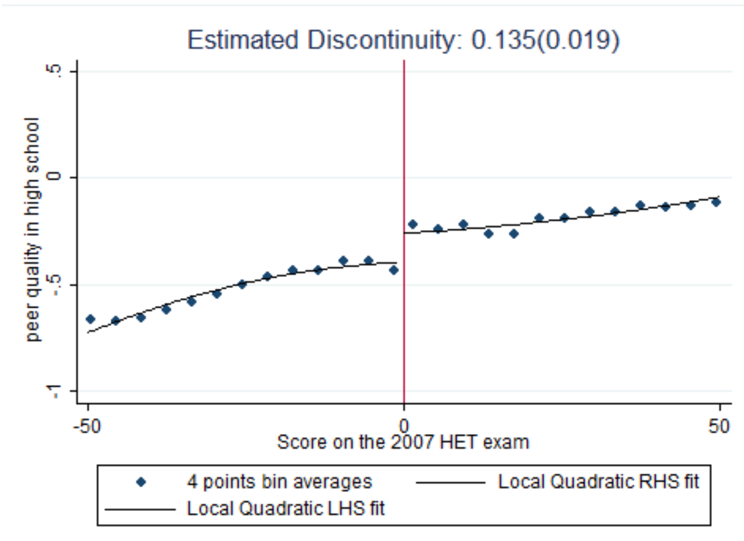
III.5.1 Effects of School Quality Across All Admission Thresholds

We begin by examining the impact of attending better schools using all of the admission thresholds in our data. Specifically, we seek to document that threshold-crossing is associated with increases in peer quality, and then ask whether threshold-crossing leads to improved performance on the high school exit exam.

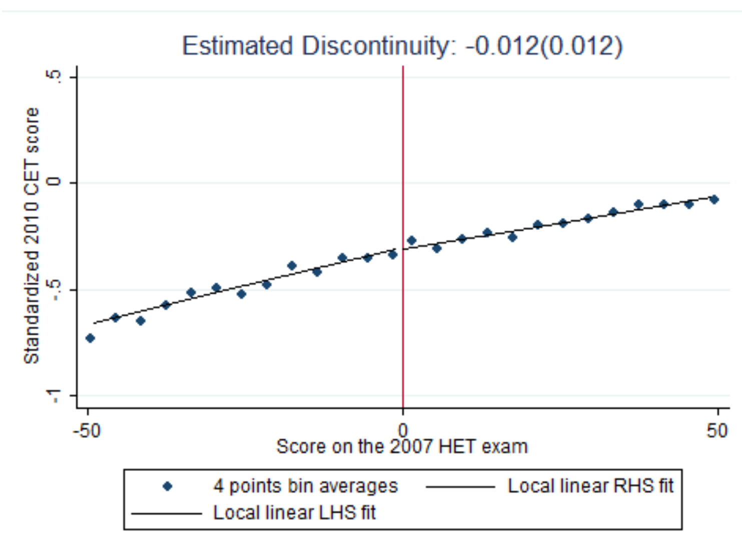
The results are shown graphically in Figure III.2. These figures take the same form as those after them in that open circles represent local averages of the outcome over a 4 point score range. We show results using a bandwidth of 50 points on either side of the cutoff. The running variable is defined as the number of points above the admission threshold. Consequently, a value of zero on the x-axis implies that the student barely met the admission threshold for the school.

Figure III.2A shows results for peer quality, defined as the average high school entrance exam score of students in the school in which the student enrolled. Using a local quadratic fit and a bandwidth of 50, we find that average peer quality significantly increases at the threshold. Specifically, threshold crossing leads to an improvement in peer quality of 13.5 percent of a standard deviation. Thus, there appears to be compelling visual evidence that threshold- crossing does lead students to attend ”better” schools, with higher-performing peers. While this relationship is deterministic given the way in which admission decisions are made, it does reflect that given the opportunity, on average students choose to enroll in schools with higher-achieving peers.

Figure III.2 Local Polynomial "Stacked RD" Estimates for Attending Better Schools

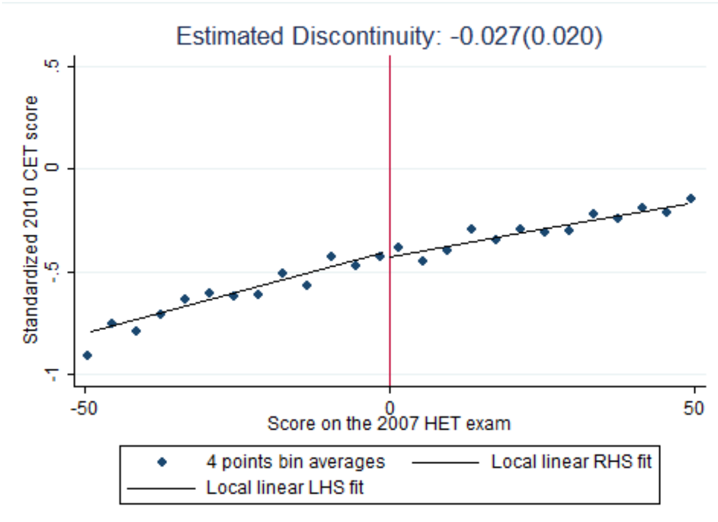


(a) Peer quality based on scores on high school entrance exam



(b) College entrance exam test scores

Figure III.2 Continued



(c) Eligibility to attend a four-year college

Notes: Sample includes students who took the high school entrance exam (HET) in the year 2007. Since we observe individuals with multiple cutoffs, we cluster at the student ID level.

Corresponding regression estimates are shown in Panel A of Table III.2. Results are shown using bandwidths ranging from three-quarters of the optimal bandwidth to 2.5 times that optimal bandwidth, where optimal CCT bandwidth was calculated as 14 points. Estimates are also shown with and without controls. Estimates for the effect of threshold-crossing on peer quality range from 0.15 to 0.2 standard deviations; all estimates are statistically significant at the one percent level.

Figure III.2B shows results for the main outcome of interest, the college entrance test score. This score is far and away the main determinant of whether a student is able to attend a four-year college, and how selective that four-year college will be. Results indicate that even though students barely above the cutoff attend significantly better schools, they do not achieve at higher levels as a result. Estimates across a range of bandwidths and specifications in Panel B of Table III.2 range from -0.017 to 0.006 of a standard deviation in CET scores. None of the estimates are statistically significant at conventional levels.

However, one might be concerned that average scores may not reflect benefits to attending better schools if students and teachers are aiming to improve scores primarily over one part of the distribution. Since an important goal of many students is to earn a score high enough to gain entry into a four-year college, we focus on an outcome that measures whether the college entrance exam score achieved exceeded the cutoff for attending four-year college in the province. Results are shown in Figure III.2C and indicate that attending better schools does not lead to improved access to four-year colleges. Corresponding estimates in Panel C of Table III.2 are similar. In short, there is little evidence that attending better schools improves cognitive ability or college attendance, on average. In addition, we also test for heterogeneity by gender. Results are shown in Appendix Figure A2, and indicate that while peer quality across the threshold is higher for both boys and girls, neither group experiences a cognitive return or increase in college attendance.

Table III.3 RD Estimates for Attending Tier I Schools

Bandwidth	2.5 CCT	2 CCT	1.5 CCT	1.25 CCT	CCT	0.75 CCT
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Discontinuity in probability of attending						
Tier I school	.631***	.632***	.671***	.649***	.642***	.628***
	(.031)	(.036)	(.046)	(.052)	(.064)	(.079)
With Controls	.637***	.637***	.677***	.656***	.654***	.637***
	(.031)	(.037)	(.047)	(.053)	(.064)	(.078)
Observations	7,167	6,046	4,654	3,901	3,133	2,389
Panel B: Discontinuity in						
high school peer quality	.301***	.314***	.372***	.357***	.358***	.345***
	(.025)	(.028)	(.032)	(.035)	(.043)	(.051)
With Controls	.315***	.329***	.395***	.378***	.380***	.359***
	(.025)	(.029)	(.033)	(.037)	(.044)	(.051)
Observations	6,680	5,578	4,278	3,642	2,886	2,237
Panel C: Discontinuity in						
CET exam scores	.082***	.089***	.083***	.085***	.073***	.069**
	(.017)	(.019)	(.022)	(.023)	(.025)	(.028)
With Controls	.081***	.086***	.084***	.090***	.073***	.065**
	(.016)	(.017)	(.021)	(.021)	(.022)	(.027)
Observations	9,909	8,870	7,352	6,454	5,298	4,112
Panel D: Discontinuity in likelihood of enrolling in						
4-year college	.135***	.097***	.063**	.049	.051	.056
	(.023)	(.024)	(.028)	(.030)	(.035)	(.043)
With Controls	.136***	.101***	.066**	.048	.048	.053
	(.023)	(.024)	(.028)	(.030)	(.034)	(.041)
Observations	8,478	7,236	5,824	4,859	3,941	3,030
Score Polynomial	One	One	One	One	One	One

Notes: Sample includes students who took the college entrance exam in the year 2007. Controls include: Age, gender, district fixed effects and middle school fixed effects. Optimal Bandwidth selected using the CCT bandwidth selector proposed in Calonico et al. (2015). Optimal BW = 20 for likelihood of attending Tier I school regressions. Optimal BW = 18 for high school peer quality regressions. Optimal BW = 34 for CET exam regressions. Optimal BW = 25 for likelihood of enrolling in four year degree regressions.
Standard errors clustered at the score level. *** p <0.01 ** p <0.05 * p <0.1

III.5.2 Effects of Attending Elite Tier I Schools

Given that much of the recent literature has focused on the returns to attending elite high schools (e.g., Abdulkadiroglu, Angrist, and Pathak, 2014; Dobbie and Fryer, 2014), we now turn to examining the returns to attending elite schools in our sample. Specifically, we examine the returns to attending Tier I, relative to Tier II schools. This cutoff is one of several used to identify effects in the previous section.

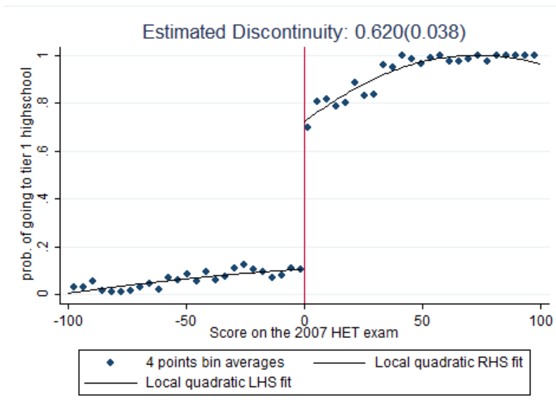
Results are shown graphically in Figure III.3. Panel A shows the likelihood of attending a Tier I high school for those just above and just below the admission threshold. Results indicate that while only around 10 percent of applicants just below the cutoff attend Tier I schools, more than 70 percent of those just above the cutoff attend Tier I schools. We note that the likely reason some students (i.e., noncompliers) are able to attend despite missing the cutoff is due to the high price admission slots allocated by the schools, as well as exceptions to the admission policy granted to some applicants such as athletes. Corresponding local linear estimates in Panel A of Table III.3 range from 63 to 67 percentage points; all estimates are significant at the 1 percent level.

Panel B of Figure III.3 shows that threshold-crossing leads to significant increases in peer ability of approximately 0.29 standard deviations of the college entrance exam. This reflects that on average Tier I schools are attended by significantly higher ability peers, though the schools may also be better in other ways (we return to this issue later). Corresponding regression estimates shown in Panel B of Table III.3 range from 0.30 to 0.37 standard deviations, all of which are significant at the 1 percent level. Thus, our results indicate that being eligible to attend a Tier I school results in roughly a 66 percentage point increase in the likelihood of attending a Tier I school, and an increase in peer quality of one-third of a standard deviation. Importantly, estimates for both the likelihood of attending Tier I schools and peer quality are nearly unchanged when adding controls measuring age, gender, and district and middle school fixed effects.

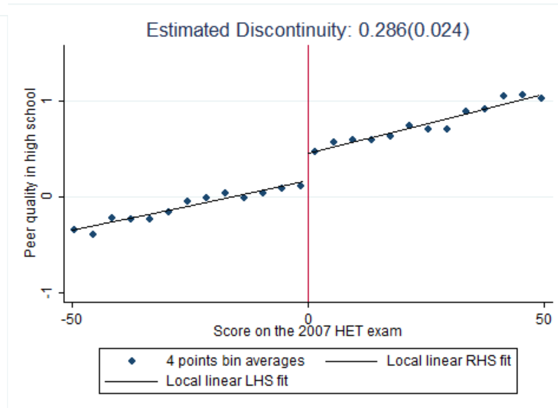
Panel C of Figure III.3 shows that in contrast to the results across all admission thresholds, being eligible to attend an elite Tier I high school leads to a nearly one-tenth standard deviation increase in achievement on the college entrance test. Corresponding regression estimates are shown in Panel C of Table III.3. The smallest estimate is that for the narrowest bandwidth (0.75 of the optimal bandwidth), which is 0.07 standard deviations and is significant at the 5 percent level. Estimates for bandwidths between 1 and 2.5 times optimal bandwidth range from 0.07 to 0.09 standard deviations and are all significant at the 1 percent level. The addition of controls does not affect estimates in a meaningful way, consistent with the identifying assumption.

Panel D of Figure III.3 shows that this increase in the college entrance test scores also results in increased eligibility to attend four-year colleges. Estimates in Panel D of Table III.3 range from 5 to 14 percentage points, though only estimates for larger bandwidths are statistically significant at conventional levels.

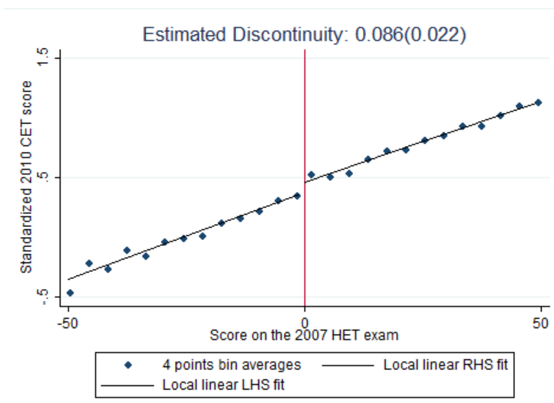
Figure III.3 Local Polynomial RD Estimates for Attending Tier I Schools



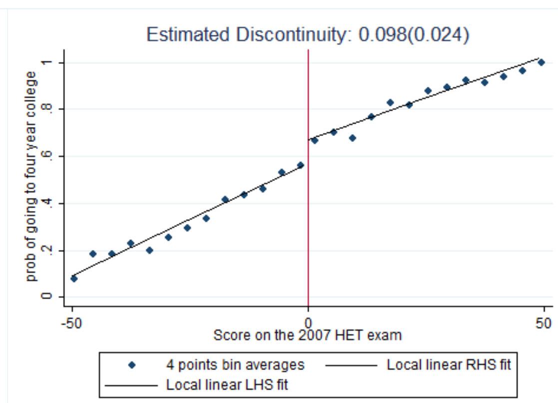
(a) Probability of attending Tier I high school



(b) Peer quality based on scores on high school entrance exam



(c) College entrance exam scores



(d) Eligibility to attend a four year college

Notes: Sample includes students who took the high school entrance exam in the year 2007. Standard errors clustered at score level.

While we do not have student-level data on long-run outcomes such as college attendance, aggregate data suggests that eligible students enroll at four-year colleges at high rates. Specifically, data from our province indicate that 63 percent of students who exceeded the eligibility threshold enrolled at four-year colleges in the province. This has important implications for long-term outcomes; Giles, Park, and Wang (2015) estimate a 37 percent return to attending four year college in China. Similarly, Li, Liu, Ma, and Zhang (2005) estimate the per-year return to attending college is as high as 10 percent. Consequently, while we lack the data to estimate the long-term returns directly, the existing literature suggests that the long-run gains to attending Tier I schools are significant.

In addition, in Figure III.4 we investigate whether returns to attending Tier I schools are different for boys than for girls. The results are quite striking; while being barely eligible for Tier I schools leads both boys and girls to attend schools with significantly higher-performing peers, only boys experience a cognitive return. Unfortunately, it is difficult for us to ascertain why this is, though we return to the issue in the next section when we discuss explanations for the overall pattern of results.

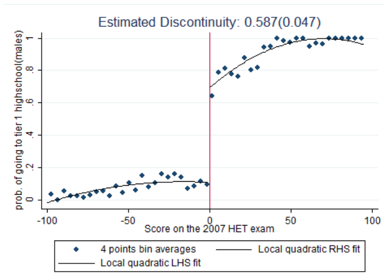
Finally, we can also report local average treatment effects of attending Tier I schools by rescaling the intent-to-treat estimates by the estimated discontinuity in the likelihood of attending a Tier I school across the admission threshold. Results are shown in Table III.4. Both males and females are roughly 60 to 65 percentage points more likely to attend a Tier I school if they are (barely) across the threshold. This speaks to the strong revealed preference for attending more selective schools in this context. In addition, Panel B of Table III.4 shows both intent-to-treat and local average treatment estimates of the difference in peer quality across the cutoff. Specifically, we estimate that attending Tier I schools results in an increase in peer quality of 0.48 standard deviations for boys and girls. Similarly, results indicate that Tier I school attendance leads to a 0.3 standard deviation increase in CET scores for boys, and increases their eligibility to attend a four-year college by 19 percentage points, with the effects remaining statistically insignificant for girls.

In summary, our analysis yields two findings. First, across all admission thresholds in the province, while being barely admitted results in enrollment at better schools with significantly better peers, it does not result in cognitive returns, on average. Second, admission at Tier I schools leads students to attend significantly better schools, which does improve cognitive outcomes, a return driven by boys. Importantly, we can reject the null hypothesis that these effects are equal.[4] These two apparently contradictory findings present a puzzle similar to that in the existing literature, which has documented mixed findings with respect to returns to high school quality. Thus, while the next section tests the robustness of these findings, after that we return to the question of why there are returns to Tier I high schools in China, but not to other "better" schools.

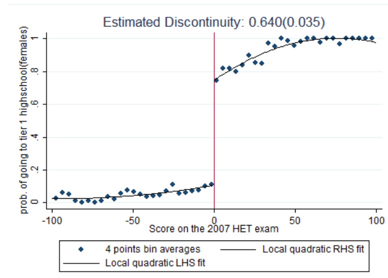
[4] For example, the estimate using optimal bandwidth with controls in column 5 of Table III.2 indicates that a 0.187 standard deviation increase in peer quality is associated with a -0.014 standard deviation increase in CET scores. Rescaling these estimates (and standard errors) by a factor of two results in a change in peer quality roughly equivalent to the 0.380 increase across the Tier I threshold shown in Panel B of Table III.3. However, the rescaled upper bound of the 95 percent confidence interval for the estimate in Panel B of Table III.2 is only 0.047, which is considerably lower than the estimate of 0.073 in column 5 of Panel B in Table III.3.

Figure III.4 Local Polynomial RD Estimates for Attending Tier I Schools, by Gender

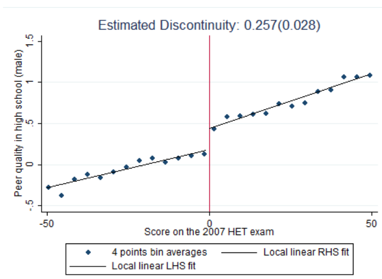
Males | Females



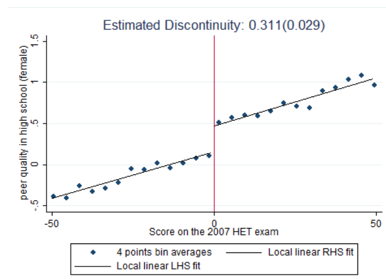
(a) Probability of attending Tier I high school



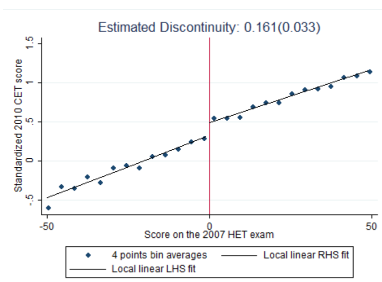
(b) Probability of attending Tier I high schoo



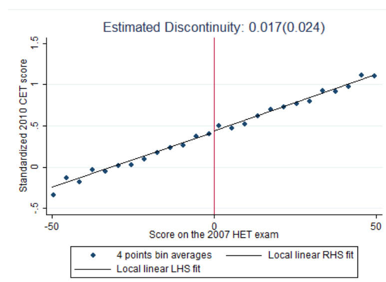
(c) Peer quality in high school



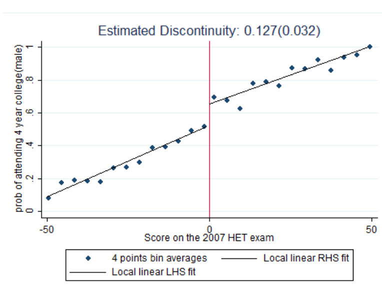
(d) Peer quality in high school



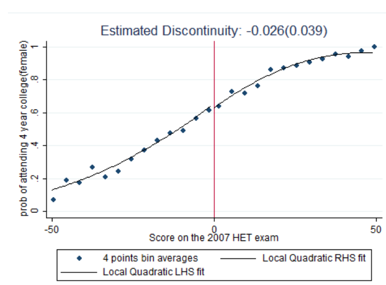
(e) CET exam scores



(f) CET exam scores



(g) Likelihood of enrolling in 4-year college



(h) Likelihood of enrolling in 4-year college

Notes: Sample includes students who took the high school entrance exam in the year 2007. Standard errors clustered at score level.

Table III.4 Local Linear Intent-To-Treat and Local Average Treatment Effect Estimates for Attending Tier I Schools

Treatment effect	ITT	LATE	ITT	LATE	ITT	LATE
Gender	All		Males		Females	
Panel A: First stage						
Likelihood of attending Tier I school	.632***	—-	.609***	—-	.652***	—-
	(.036)		(.045)		(.033)	
Panel B: Discontinuity in school inputs						
High school peer quality	.303***	.486***	.290***	.484***	.317***	.489***
	(.026)	(.023)	(.030)	(.029)	(.032)	(.035)
Panel C: Discontinuity in outcomes						
College entrance exam test scores	.094***	.155***	.174***	.306***	.015	.020
	(.024)	(.040)	(.034)	(.061)	(.027)	(.042)
Eligibility to attend a 4-year college	.071***	.119***	.103***	.191***	−.042	−.060
	(.027)	(.043)	(.036)	(.059)	(.042)	(.063)
Observations	6046	6046	2813	2813	3233	3233

Notes: Sample includes students who took the high school entrance exam in the year 2007. All regressions include controls: gender, age, district fixed effects and junior high school fixed effects. For ease of comparison, all local linear regressions use an equal bandwidth of 40 points on either side of the cutoff. Standard errors are clustered at the score level. *** p <0.01 ** p <0.05 * p <0.1

III.5.3 Threats to Identification

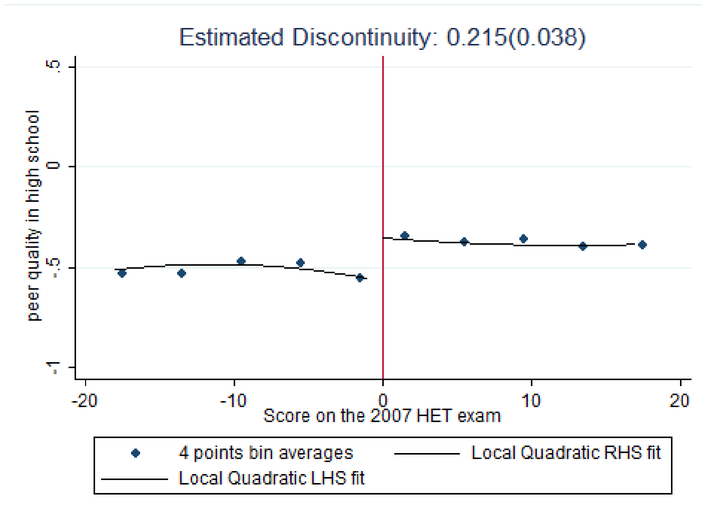
One potential threat to identification is if attending better schools leads students to select a different academic track, a decision that is made in the second year of high school. For example, if (barely) going to a Tier I high school increases the likelihood of a student choosing a scientific track, then that difference, rather than a broader sense of improved school quality, could drive our results. To test for this, we check whether the probability of choosing an arts versus science track is discontinuous at the threshold for attending a Tier I high school. Results are shown in Appendix Figure A5a and A5b. Both figures show that the likelihood of majoring in arts versus science is smooth across all admission thresholds (Figure A5a) as well as the cutoff for Tier I schools (Figure A5b).

In addition, we also test whether differential grade repetition across the admission cutoff could bias our estimates. For example, if Tier I high schools were more likely to have their worst students repeat a grade, then perhaps the improvement in CET scores we document could be due to age or quantity of schooling, rather than school quality. While grade repetition is uncommon in China, we test explicitly for this explanation by examining whether exact age at the time of taking the CET is smooth across the admission threshold.[1] Results are shown in Appendix Figures A5c and A5d, which show that age is smooth across both cutoffs, indicating that grade repetition is unlikely to explain our findings.

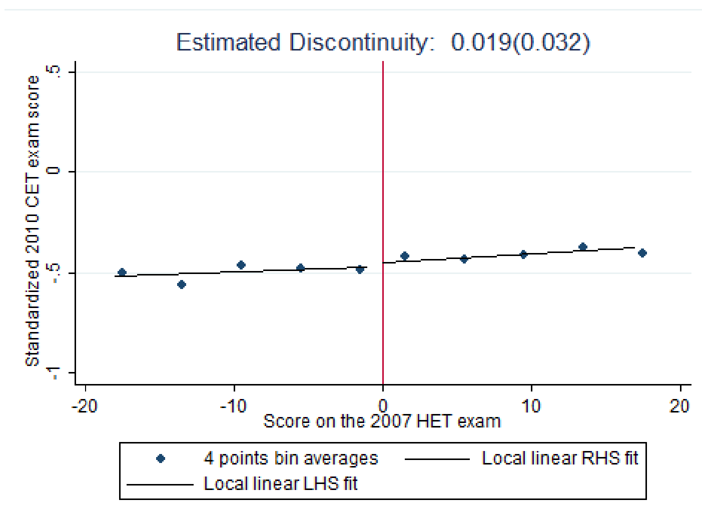
---

[1] Grade repetition is rare in part because students are not allowed to repeat their senior year of high school. For other years, in order to repeat a year a student must fail three classes after taking a make-up exam and must gain the approval of school and city-level administrators.

Figure III.5 Local Polynomial "Stacked RD" Estimates With all Cutoffs Except the Tier I Cutoff



(a) High School Peer Quality

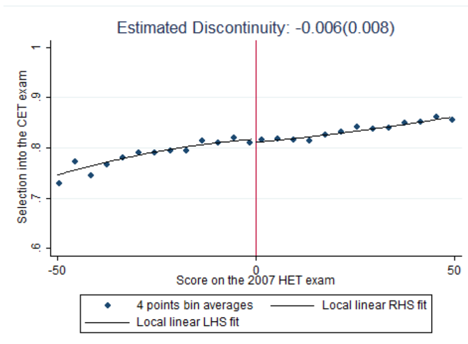


(b) Standardized CET scores

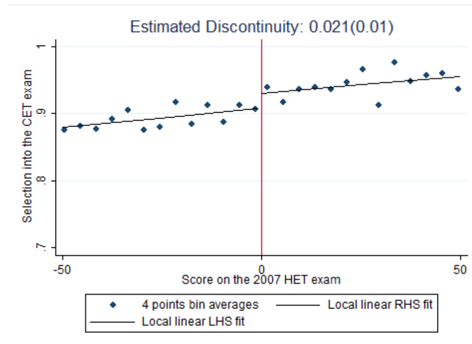
Notes: Sample includes students who took the high school entrance exam (HET) in the year 2007. Since we observe individuals with multiple cutoffs, we cluster at the student ID level. In order to exclude the Tier I admission threshold, we use a maximum bandwidth of 18 points.

Figure III.6 Selection into the College Entrance Exam

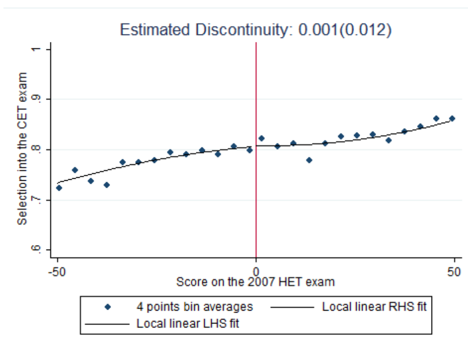
Attending Better Schools (All Cutoffs) | Attending a Tier I School



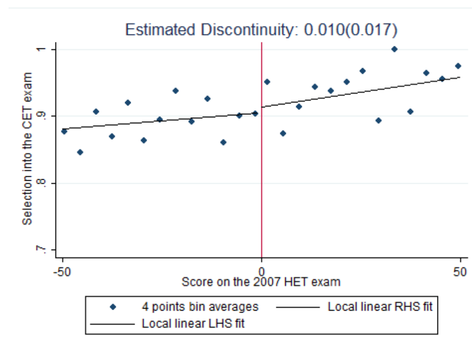
(a) Selection into the CET exam



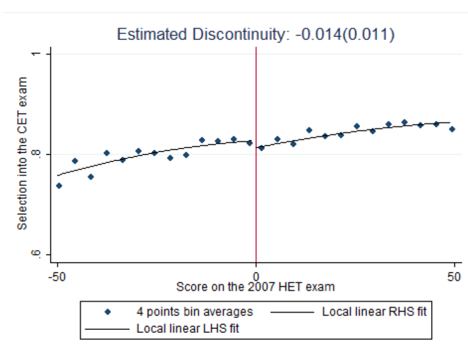
(b) Selection into the CET exam



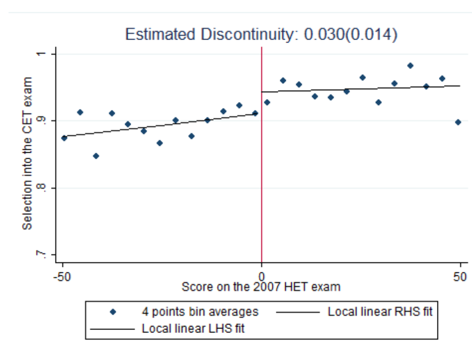
(c) Selection into the CET exam (Males only)



(d) Selection into the CET exam (Males only)



(e) Selection into the CET exam (Females only)



(f) Selection into the CET exam (Females only)

Notes: Sample includes students who took the high school entrance exam in the year 2007 (including those with no college entrance exam scores).

Table III.5 Regression Discontinuity Estimates for Selection into the College Entrance Exam

Bandwidth	2.5 CCT	2 CCT	1.5 CCT	1.25 CCT	CCT	0.75 CCT
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: (Going to a better school)						
Selecting into the CET entrance exam (All)	−.011	−.007	−.001	.004	−.000	.010
	(.008)	(.008)	(.011)	(.012)	(.013)	(.014)
Females only	−.015	−.017	−.012	−.008	−.001	−.003
	(.009)	(.011)	(.014)	(.016)	(.017)	(.020)
Males only	−.004	.004	.010	.017	.001	.025
	(.010)	(.012)	(.016)	(.017)	(.019)	(.021)
Observations (females)	18226	14615	10944	9076	7461	5552
Observations (males)	16522	13177	9840	8246	6768	5082
Panel B: (Going to a top school)						
Selecting into the CET entrance exam (All)	.007	.020**	.020**	.017	.015	.025*
	(.008)	(.009)	(.010)	(.011)	(.012)	(.014)
Females only	.017	.037***	.021	.019	.011	.025
	(.013)	(.014)	(.015)	(.017)	(.019)	(.022)
Males only	−.005	.001	.019	.015	.019	.023
	(.015)	(.017)	(.019)	(.021)	(.024)	(.027)
Observations (females)	5386	4663	3791	3214	2607	1983
Observations (males)	4800	4166	3375	2862	2363	1822
Score Polynomial	One	One	One	One	One	One

Notes: Sample includes students who took the high school entrance exam in the year 2007 with known high school cutoffs (including those who did not site for the 2010 college entrance exam). Optimal Bandwidth selected using the CCT bandwidth selector proposed in Calonico et al. (2015). Optimal BW = 12 for likelihood of taking CET exam (going to a better school).
Optimal BW = 29 for likelihood of taking CET exam (going to a top school). *** p <0.01 ** p <0.05 * p <0.01.

Table III.6 Bounding Analysis for the Estimated Impact of Attending Tier I Schools

Bandwidth	84 points (1)	68 points (2)	51 points (3)	43 points (4)	34 points (5)	26 points (6)
Panel A: Selection into the CET exam	.003	.009	.020**	.017*	.013	.017
	(.009)	(.009)	(.009)	(.010)	(.011)	(.012)
Constant (Control mean)	.926	.916	.903	.904	.908	.902
Proportion to be trimmed	0.32%	0.98%	2.2%	1.88%	1.43%	1.88%
Panel B: College entrance exam test scores (Original regression estimates)	.082***	.089***	.083***	.085***	.073***	.069**
	(.017)	(.019)	(.022)	(.023)	(.025)	(.028)
College entrance exam test scores (Lower bound estimates)	.083***	.095***	.085***	.080***	.070**	.062**
	(.018)	(.018)	(.024)	(.023)	(.028)	(.030)
College entrance exam test scores (Upper bound estimates)	.091***	.126***	.149***	.139***	.112***	.104***
	(.020)	(.020)	(.022)	(.024)	(.028)	(.030)
Observations	9909	8870	7352	6454	5298	4112
Observations after trimming	9851	8826	7268	6391	5258	4068
Score Polynomial	One	One	One	One	One	One

Notes: Sample includes students who took the high school entrance exam in the year 2007 with known high school cutoffs (including those who did not site for the 2010 college entrance exam). Optimal Bandwidth selected using the CCT bandwidth selector proposed in Calonico et al. (2015). To ease comparison with our previous estimates, we use the same bandwidths predicted by the CCT for the original college entrance score regressions. Bootstrapped standard errors reported for the upper and lower bound estimates *** p <0.01 ** p <0.05 * p <0.1.

Perhaps a more worrisome potential source of bias is the possibility of selection into taking the college entrance exam. That is, if barely being admitted to a better school (or an elite school) made it more or less likely for the student to take the college entrance test, then our estimates could be biased. We address this concern by first testing for selection into test- taking, and then using those estimates to bound our estimates. We begin by matching our dataset to data containing the population of high school test takers, regardless of whether they took the high school entrance exam. In this way, students who do not match are identified as not having taken the CET. Our match rate is high; we estimate that 91 percent of students entering high school sat for the CET exam. This is in line with official aggregate data for the districts in our sample, which indicate that 90 to 94 percent of students take the CET exam over the years we study.

We then examine whether going to a better quality high school leads students to take the CET at different rates. Results are shown in Figure III.6. Results indicate that while going to any better school is not associated with differential college entrance test-taking, barely attending an elite high school does appear to lead to a higher rate of test-taking. Additional results in Figure III.6f indicate that this increased test-taking is driven by girls; in contrast, the rate of test-taking is constant across the admission threshold for boys (Figure III.6d).[1] Corresponding estimates in Table III.5 yield the same conclusion: while there is no evidence of selection into test-taking for boys and girls when looking at better schools overall, students who are barely admitted to elite schools are one to two

---

[1] This could also potentially explain results from the previous section indicating that the likelihood of observing a female in the sample is higher at the Tier I cutoff for some bandwidths. See Appendix Table A1.

percentage points more likely to take the CET, though these estimates are not statistically significant across all bandwidths.

The simplest explanation for this finding is that some marginal students are induced to take the CET when barely attending Tier I schools, when they would not have if they had attended lower-tier schools. This would likely work against our finding that attending Tier I schools leads to improved CET performance. In addition, we note that the positive returns to attending Tier I schools were driven by boys, while Figure III.6d shows no evidence of selection into test-taking for boys. This provides additional comfort that the type of selection into test-taking that we observe cannot explain our findings.

Nevertheless, we perform formal bounding exercises to assess the degree to which this selection into test-taking could affect our results. Specifically, we use the trimming procedure suggested by Lee (2009). The intuition behind this test is as follows. To find a lower bound (worst case scenario) for the estimated impact of treatment on college exam scores, we assume that only the best students attending the most selective high schools, who would have otherwise dropped out, select into the exam. Thus, dropping the top distribution of the treatment group makes it comparable to the control group. Formally, the share of students to be trimmed from the treatment group is derived from the treatment estimate of the likelihood of selecting into the college entrance exam.[2] A similar procedure—trimming the bottom performing students in the treatment group—is used to estimate the upper bound.

Table III.6 summarizes the updated local linear regressions by comparing previous college test score RD estimates with those estimated using the trimming analysis. Bootstrapped standard errors are reported in parentheses for the lower and upper bound estimates. For consistency and comparability, we use the same bandwidths as the college entrance exam score regressions in Table III.3. We present lower and upper bound estimates for all bandwidths. Results indicate that the lower and upper bound estimates of the return to attending an elite high school remain positive and significant, and range from 6.2 to 15 percentage points, compared to original estimates ranging from 6.1 to 9.4 percentage points. For example, bounds corresponding to the estimate using optimal bandwidth are shown in Column 5. Our primary estimate was 7.3 percentage points, while the estimated lower and upper bounds are 7 and 11.2 percentage points, respectively, both of which are statistically significant at the 5 percent level. Thus, we conclude that selection into the CET does not bias our results in a meaningful way.

[2] For example, using a local linear regression of bandwidth =50, we estimate that students are 1.69 percentage points more likely to select into the college entrance exam at the cutoff. To estimate the total percent of students to be trimmed, we merely divide 1.69 by the mean proportion of test takers for the control group at the threshold (90.5%). This results in us trimming 1.85 percent of the data.

III.6 Interpretation: The Role of Peer Quality, Class Size, and Teacher Quality

Perhaps the simplest potential explanation for why there are returns to quality for Tier I high schools, but not for others, is that there are smaller or nonexistent differences in peer quality across the other cutoffs that are obscured when the results are aggregated, as in Figure III.2. To examine whether or not that is the case, we ask whether there is a significant difference in peer quality across cutoffs other than the Tier I cutoff. Results are shown in Figure III.5, which stacks together all cutoffs other than the Tier I cutoff.[3] Results in Figure III.5a indicate that while barely admitted students attend schools with students who scored one- quarter of a standard deviation higher on the high school entrance exam, Figure III.5b indicates that they score no higher on the college entrance exam. Thus, it is clear that while peer quality does increase discontinuously across the non-Tier I admission cutoffs, it is equally clear that there is no evidence of improvement on the CET.

Results for each cutoff separately are shown in Appendix Figure A3. While splitting the sample in this way leads to reduced statistical power, results are consistent across all three different sets of admission thresholds in showing significant increases in peer quality but no evidence of return. Panel (a) shows the discontinuity in peer quality at the Tier II cutoff of about one-fifth of a standard deviation, while panel (b) shows that there is no evidence of performance gains to barely attending the better school. Similarly, panel (b) shows the discontinuity in peer quality of 0.4 standard deviations

[3] Because we exclude the Tier I cutoff, we cannot use a bandwidth greater than 18 points. This is because for one of our districts, the Tier I cutoff is 634 points, while the next cutoff after that is 615 points.

across the cutoffs within Tier I (i.e., attending a more selective Tier I school versus a less selective Tier I school), while panel (c) shows that there is no performance gain across that cutoff. Finally, panel (e) documents a 0.11 standard deviation increase in peer quality across the cutoffs within Tier II (i.e., more selective versus less selective schools within Tier II), while panel (f) reveals no positive cognitive return for this group of students. Thus, Figure A3 shows that while there are significant improvements in peer quality across all non-Tier I admission thresholds in the school quality distribution, there are no cognitive returns to attending the more selective schools.

Combined with our main results reported earlier, these findings indicate that while barely attending Tier I schools (and the 0.35 standard deviation increase in peer quality) results in an improvement of 7 to 9 percent of a standard deviation in CET scores, barely attending better schools across all other cutoffs except the cutoff from Tier II to Tier I (and the 0.2 standard deviation average increase in peer quality) does not result in better CET scores. Thus, these results suggest that peer quality is unlikely to explain the differences in cognitive returns across the Tier I and non-Tier I high schools that we observe.

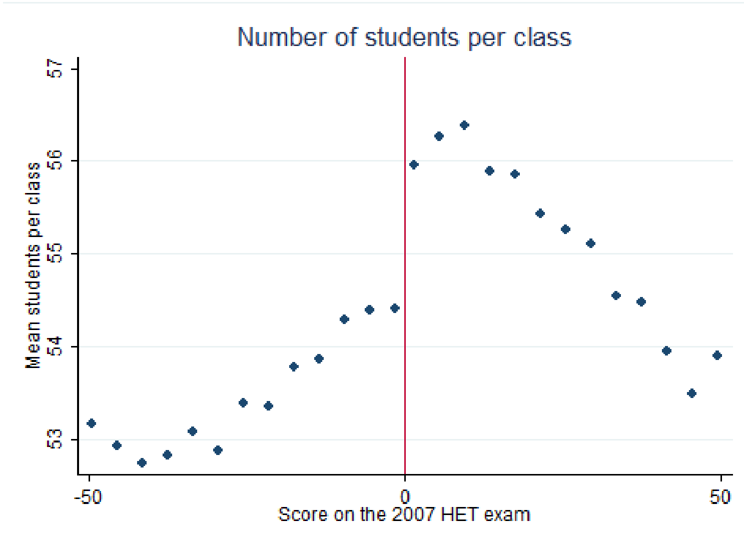
A more nuanced explanation, however, is that perhaps nonlinear returns to peer quality could explain the observed pattern of findings. For example, one might argue that higher peer quality benefits high-ability students more than low-ability students. We also view that as inconsistent with the evidence. Specifically, we note that even within Tier I—where all students are relatively high-ability—we find a significant increase in peer quality without observing an increase in cognitive performance. Specifically, Figure A3c shows that those students who barely attend better Tier I schools versus worse Tier I schools experience peer quality that is 0.4 standard deviations higher, while Figure A3d shows that these students do not perform better on the CET exam.

A related possibility is that perhaps students sort differently into peer groups across the different cutoffs. For example, if the barely admitted students at the non-Tier I cutoffs were to primarily associate with lower-performing students at those better schools, but the barely admitted students at the Tier I cutoff were to associate with higher-performing students (i.e., mix better, or more randomly), then that could explain the heterogeneity in returns. While our data do not allow us to directly test this, given the consistency of findings across cutoffs shown in Figure A3, we view this explanation as implausible. As a result, we interpret the pattern of findings as inconsistent with the hypothesis that the returns to school quality in this context are driven by peer quality.

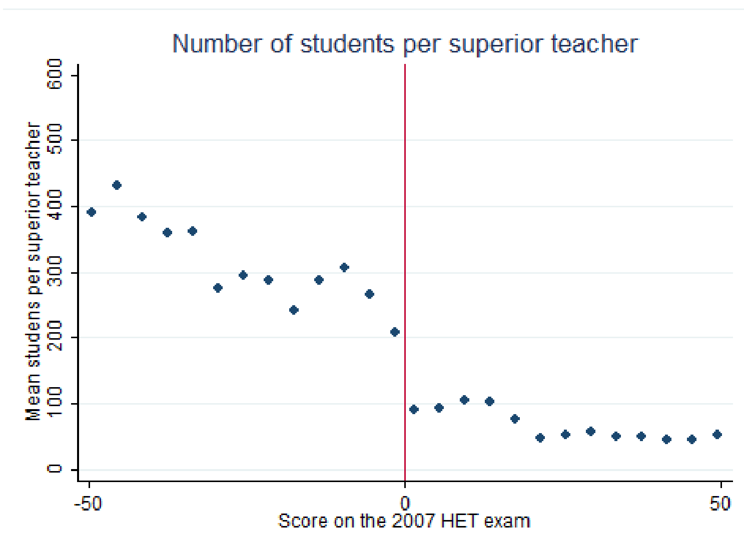
A second potential interpretation is that the difference in returns is due to differential behavioral responses by students, such as those documented by Pop-Eleches and Urquiola (2013) in Romania. For example, in response to attending better schools students could feel marginalized by being the worst students relative to others, or may receive less parental effort as a result of attending the better school. While we are unable to test directly for these behavioral responses, we view them as unlikely to explain the heterogeneity in findings. That is because it is difficult to imagine how parents (or students) react differently across the Tier I cutoff than they do across cutoffs within Tier I, within Tier II, or across the Tier II/Tier III cutoff. Thus, while we cannot rule out this type of differential-response explanation directly, we view it as unlikely to explain the results.

Next, we turn to whether the heterogeneity in returns could be driven by differences in class size, which has received considerable attention in the literature (e.g., Angrist and Lavy, 1999; Krueger, 2003). For example, if class size is discontinuously smaller across the Tier I cutoff, but not across other cutoffs, it could explain the heterogeneity in returns that we observe. Results for the Tier I cutoff are shown in Figure III.7a, and indicate that students who barely attend Tier I schools are in significantly larger classes (56 versus 54 students). In contrast, Figure III.8a shows that across all other cutoffs, class size is no different. If anything, that suggests that return to attending Tier I schools should be lower than the return across other cutoffs. As a result, the observed heterogeneity of returns is not easily explained by differences in class size.

Figure III.7 Class Size and Teacher Quality RD Estimates for Attending a Tier I School



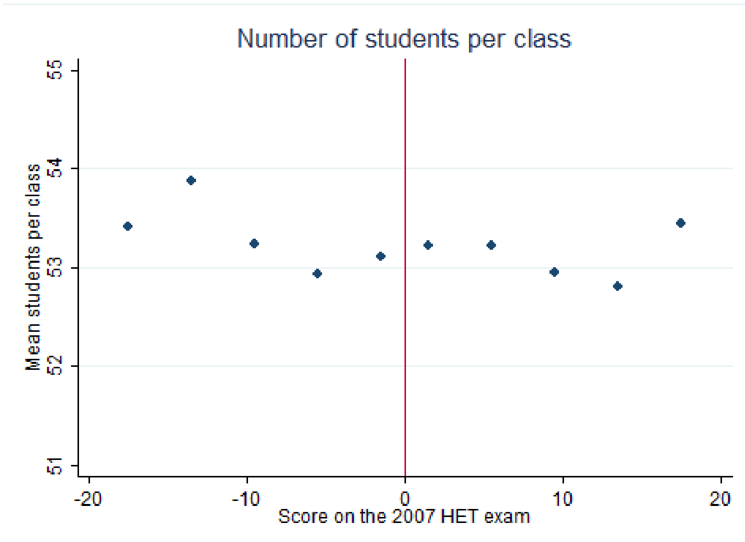
(a) Number of students per class



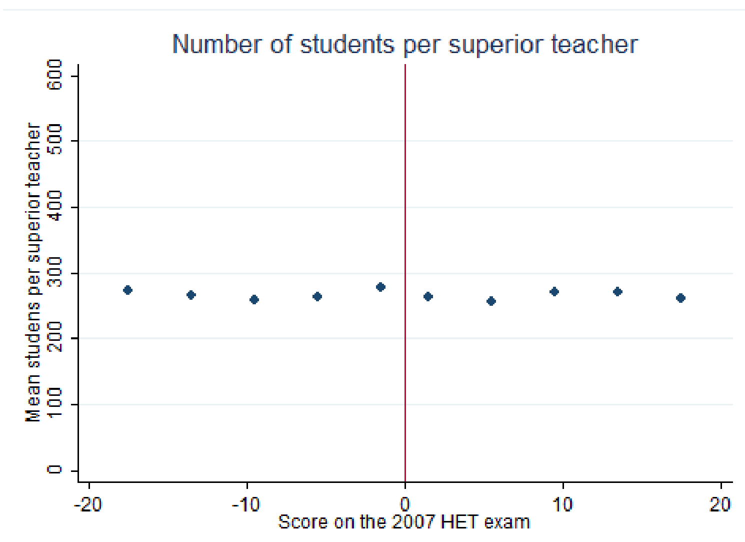
(b) Number of students per superior teacher

Notes: Sample based off of school level data.

Figure III.8 Class Size and Teacher Quality "Stacked RD" Estimates for All Admission Cutoffs Except for the Tier I Cutoff



(a) Number of students per class



(b) Number of students per superior teacher

Notes: Sample based off of school level data. In order to exclude the Tier I admission threshold, we use a maximum bandwidth of 18 points.

Another potential explanation is teacher quality, which has been shown to lead to significant increases in achievement in other settings. While we do not have the necessary data to estimate teacher value-added in our setting, we do observe the proportion of teachers with the "superior teacher" ranking, which is the top ranking a teacher can receive in this province, and the only ranking one cannot automatically qualify for with tenure and advanced degrees. In addition, Lai, Sadoulet, and de Janvry (2011) use lottery data from Beijing middle schools to show that teacher ranks are highly correlated with the estimated school fixed effects. Similarly, Hannum and Park (2001) report that teachers with the highest rank in their sample increase achievement by 0.29 standard deviations relative to teachers with the next highest rank, and conclude that the quality ranks used in the Chinese schooling system contain a substantial amount of information on teacher quality that is not contained in conventional measures such as education of the teacher and years of experience.

Results are shown in Figures III.7b and III.8b, and indicate that the only discontinuity in teacher quality is at the Tier I admission threshold, and not at the non-Tier I thresholds. Appendix figures A4d, A4e, and A4f further break down the non-Tier I thresholds and show no evidence of discontinuities in teacher quality at Tier II versus III, within Tier I, or within Tier II. This pattern is thus broadly consistent with our results on cognitive returns shown above, and suggests that returns to high school quality in this context are due to teacher quality, rather than to peer quality. This is also broadly consistent with work by Jackson (2013), who reports that peer achievement can only explain a small fraction of the school selectivity effect in Trinidad & Tobago.[4]

To assess whether the differential exposure to superior teachers can explain all of the improved achievement we observe, we perform a back-of-the-envelope calculation. We start with our estimate from Panel A of Figure III.6, which indicates that barely admitted students attend schools where the ratio of students to superior teachers increases from approximately 200/1 to 100/1. Given our findings on the return to attending Tier I schools shown in Panel C of Table III.3, we estimate that if superior teachers increased achievement by 0.13 to 0.17 standard deviations relative to their counterparts, then the additional superior teachers would explain all of the estimated return to attending Tier I schools.[5] We view such effects as plausible, if slightly larger than the finding in the literature that a one standard deviation increase in teacher quality results in roughly a one-tenth increase in achievement (Rockoff 2004; Rivkin, Hanushek, and Kain 2005; Aaronson et al. 2007; Kane, Rockoff, and Staiger 2008). In short, we

[4] What is less clear is why Tier I schools and the associated increase in teacher quality benefit primarily boys, rather than girls, as shown in Figure III.4. In results not shown, we document that there is no difference across the Tier I threshold in the proportion of male teachers; suggesting overall teacher gender does not explain the difference in performance. Thus, we view the possibility that superior teachers have disproportionate positive impacts on boys learning as the most likely explanation, though we do not have the data to provide evidence on the mechanism through which that might be the case. For example, we do not have data on the gender mix of superior teachers.

[5] The average class size at the cutoff is 36 students, which means a doubling of the superior teacher ratio implies that an additional 18 percent of students in Tier I schools are taught by superior teachers. Estimates in Panel C of Table III.3 indicate that Tier I schools increase average student achievement by 0.07 to 0.09 standard deviations. Assuming that all of that return comes from increased access to superior teachers implies rescaling those estimates by 0.18, which indicates that superior teachers increase scores by 0.38 to 0.50 standard deviations. Dividing by three given the three years of high schools results in estimates of 0.13 to 0.17.

estimate that given the increased access to superior teachers in Tier I schools, if those teachers have a value-added that is around one-and-a-half standard deviations better than their counterparts, superior teachers would explain the entire cognitive return to attending Tier I high schools.

In summary, the additional exercises in this section indicate that the returns to high school quality in our setting are unlikely to be caused by peer quality or class size. Rather, the empirical evidence suggests that the returns are likely due to differences in teacher quality - as proxied by superior teacher rank. Thus, while we cannot rule out the possibility that the heterogeneous returns across the cutoffs are due to differences in some unobserved input other than teacher quality, we think the interpretation most consistent with our pattern of findings is that differences in teacher quality are responsible for returns to school quality in this context.

III.7 Conclusion

This paper estimates the cognitive benefits due to attending higher quality high schools. It does so by using a regression discontinuity design that compares the academic outcomes of students who are barely eligible and ineligible to enter better quality high schools in China. Results indicate that across distribution of school quality, the only positive returns to school quality are for those who attend Tier I, rather than Tier II schools. Importantly, we document that this is true despite the fact that admission threshold-crossing across the continuum of high school quality is associated with significant increases in peer quality. As a result, we conclude that at least in this setting, positive cognitive returns to high school quality are unlikely to be due to peer quality.

We provide additional evidence that the returns to attending elite Tier I schools are due to teacher quality. Specifically, we document that students who (barely) attend Tier I schools are twice as likely to be taught by teachers of superior rank. A back-of-the-envelope calculation suggests this increased exposure to teachers of superior rank can explain the entire cognitive return to Tier I schools if those teachers increase achievement by around 0.15 standard deviations compared to their counterparts. Given the existing literature on the impact of teacher quality, this suggests the teacher quality of superior-rank teachers is roughly 1.5 standard deviations higher than that of their counterparts. This is plausible given existing research on education in China, which has shown that teachers of superior rank increase test scores by around 0.30 standard deviations and that teacher rank is highly correlated with estimates of school quality (Hannum and Park, 2001; Lai, Sadoulet, and de Janvry, 2011).

These findings have important implications for the literature on the returns to school quality. First, by demonstrating the heterogeneity of returns in a single educational context that cannot be explained by differences in peer quality, the results here highlight the im- portance of measuring additional education inputs. Thus, while peer quality has been and remains a straightforward proxy for school quality, the results of this paper highlight that it is not a sufficient statistic for school quality, and that focusing on peer quality can make it difficult to reconcile seemingly inconsistent findings.

In addition, to the extent that returns to high school quality more generally are not due peer quality, it has important implications for how to increase academic achievement. If the benefits to attending better schools were due to better peers, it would be difficult to extend those benefits more broadly since there is a limited set of high-performing peers. In contrast, the results presented here demonstrate that at least in this context, policymakers may be able to do other things—such as improve teacher quality—to replicate school quality and improve educational outcomes.

CHAPTER IV

IMPACTS OF HIGH PERFORMING PEERS ON FEMALE STEM CHOICE IN HIGH SCHOOL

IV.1 Introduction

The question of why there are fewer women in science, technology, engineering and math (STEM) majors and jobs has long been of concern to social scientists, educators and policy makers. While women hold 48% of all jobs and half of the college-educated jobs, they only make up 24% of STEM workforce (Beede et al., 2011). STEM employment is a critical component of a country's competitiveness and the gender wage gap is relatively smaller in STEM jobs than that in non-STEM jobs. Therefore, engaging more women into the STEM-related occupations is of importance to the nation's sustainable growth potential as well as closing the gender gap in wages.

It is reported that this gender disparity has long began since high school. Only 3 percent and 2 percent of US high school girls reported an interest in the engineering and technology field respectively, compared to 31 percent and 15 percent for boys. Of all bachelor's degrees earned by women, only 13 percent of them were in a STEM field, while 28 percent of bachelor's degrees earned by men were in STEM fields. As of graduate degrees, these numbers becomes 10 percent and 24 percent for women and men respectively ("The U.S. News/Raytheon STEM Index Shows Gender and Racial Gaps Widening in STEM Fields," 2015).

Possible explanations of gender disparities in STEM education and occupations are: difference in women and men's quantitative ability or test preparation; strong gender stereotypes that discourage women from pursuing STEM education and STEM jobs; difference in the way men and women respond to STEM education and employment incentives; lack of female STEM role models etc. (Clark Blickenstaff, 2005)

Meta analysis shows that the gender differences in Grade 2-11 mathematics have been shrinking over the years and insignificant now for the general population (Hyde, Fennema, & Lamon, 1990). The gender difference in average math scores is only 0.0065 standard deviation and even in the 95^{th} and 99^{th} percentile, the gender ratio favoring males is also fairly small for whites and is even reversed for Asian Americans (Hyde, Lindberg, Linn, Ellis, & Williams, 2008). However, the tests whose scores were used to calculate these statistics fail to assess test takers' complex problem solving skills. Gender differences in advanced science classes taking still persist in the high school level.

In the US, STEM and non-STEM paths diverge at the college level. As of 2012, more than 57% of the bachelor's degrees awarded in the US are awarded to women. Although women earn more than 50% of Science and Engineering degrees combined[1], the number of women in engineering field is still only less than 20% and there has been no sign of closing of this gap. Female ratio in computer sciences has even dropped from

---

[1] Female ratio in sciences by field: agricultural sciences (53.7%), biological sciences (59.3%), computer sciences (18.2%), earth atmospheric and ocean sciences (39.1%), mathematics and statistics (43.1%), physical sciences (40.6%), psychology (76.7%) and social sciences (54.7%).

27.5% in 2002 to 18.2% in 2012 ("Women, Minorities, and Persons with Disabilities in Science and Engineering," 2014). According to the 2015 U.S. News/Raytheon STEM Index, boys of all demographics scored at least 30 points higher on the math section of the SATs than girls. By the quantitative nature of STEM fields, relatively higher math SAT scores gives men advantages to women. However, the differences in representation in sciences and quantitative fields are disproportionately larger than the distribution of SAT math scores. Turner and Bowen (1999) estimated that the difference in SAT scores account for only half of the total gender gap in college major choices (Turner & Bowen, 1999). This gender gap is also unlikely to be of pecuniary reasons. The gender wage gap is 14% for STEM jobs, 7 percentage points lower than that of non-STEM jobs. Survey also shows that women care less about pecuniary outcomes than men when it comes to choosing college major (Zafar, 2013).

Differences in quantitative ability / test preparation (if any) and pecuniary preferences only explain part of the gender gap in the STEM fields. Other socio-psychological factors such as gender stereotyping, lack of female role model in science and engineering and hostile environment for female in science classes are also potential barriers for women to pursue the science path (Clark Blickenstaff, 2005). These factors may lower women's self-esteem and confidence in learning science and previous attempts in reducing gender achievement gap by changing instructional methods may do very little to change them. Recently the Science magazine showed that positive psychological affirmation in the classroom can effectively close the gender gap in achievement by 64% to complete elimination (Miyake et al., 2010). Miyake et al implemented a 15-minute writing exercise into an introductory physics class in college where the treatment group wrote about the how a value is most important to them and the control group selected the least important value to them and wrote about how it might be important to other people. These writing sessions were given three times before the first midterm exam at week 5 and were proved to have prolonged positive effects in raising women's achievement and women's only, reducing the gender gap from 0.93 (control group) to 0.18 (treatment group). This value affirmation was also shown to be beneficial for women who tend to believe in the gender stereotype that men outperform women in physics. Affirmation was able to eliminate the negative effect of stereotype endorsement on achievement for women. These findings may be an indication of the social-psychological effect of women feeling more assured of themselves against potential intimidating environment or social-psychological influences against females in the science field.

In line with this notion, having more female peers who are high achieving in quantitative fields may also have a similar role model or affirmation effect for a female student in that she might feel more confident about the prospect of doing well in these fields. In this paper, I examine the high performing female peer effect on female student's science path choices in a high school setting. Utilizing a unique dataset from China where high school students have to choose either science or arts path that leads to the college entrance exam, I track down females opting in to or out of the science track as early as 16 years of age. I use the math score of their high school entrance exam as the measure of "high performing" and also their own entering scores as controls of their own "ability". The outcome variable is individual students' choice of whether taking the college entrance exam in "science" or "arts". To control for unobserved characteristics of schools and students that might be correlated with high performing female peer composition, we rely on idiosyncratic variation in the proportion of high performing female students across cohorts within the same school.

The rest of the paper is organized as follows. Section II introduces the institutional background of Chinese education system and discusses the identification strategy. Section III describes the data. Section IV presents the results of the school fixed effects estimates of high performing female peer effects on high school students' science path choices and Section V concludes.

IV.2 Identification

IV.2.1 Institutional Background

Children in China generally start elementary school (1st grade) at around 6-7 years old. After 6 years of elementary school, children go on to the first part of middle school, a 3-year junior middle school (7th to 9th grade) to complete the 9-year national compulsory education. 9th grade graduates then can choose to go to a vocational school or the second part of middle school (10th to 12th grade), which is equivalent to US high schools. Vocational and traditional high schools are then followed by vocational colleges (3-year or 4-year) and traditional colleges/universities (2-year or 4-year) respectively. Traditional high schools prepare students for traditional colleges/universities and are competitive as there are only enough seats for about 60% of the junior middle school graduates in our sample.[2] In order to go to a traditional high school, 9^{th} graders compete in the High School Entrance Test (HET) held by the city. In the HET, students are tested on seven subjects including Chinese language, Mathematics, English language, Physics, Chemistry, Political Science and PE[3]. The total score of the 7 subjects is the one and only criterion of high school admission for most students[4]. As our sample contains only students on the traditional track so from now on, our discussion is limited to traditional high schools and colleges/universities.

The three-year high school is neither compulsory nor free like elementary and junior middle school phases. However in most parts of China, the majority of high schools are public schools, which is also true to other levels of education in the country. These public schools do charge tuition but the amount is generally cheap (in our sample it is around $200/year and deductible if family income qualifies certain requirements) and the same across all public schools.

For college bound students, better high schools in a large extend translate into better chance of going to a better college (or any college). The city officials assign students to schools centrally in accordance to their total scores and preferences,

---

[2] Technically vocational high school graduates can also choose to take the college entrance test. However the curriculum in vocational schools differ substantially from the materials on the test thus if a student wants to go to college, she will choose to go to a traditional high school.
[3] Chinese language, Mathematics and English language are worth 150 points each; Physics, Chemistry and Political Science are worth 100 points each; P.E. is worth 40 points but is not tested in all years in our sample.
[4] The only exceptions are students with special talent, for example athletes. These students should be of a very small portion of the whole population.

following what we refer to as the "Boston mechanism"[5]. To raise funds, public schools are allowed to assign no more than 10% of their seats as "high priced" seats where a one-time fee is collected if a student choose to enter through this channel[6]. High priced seats for a school is treated as an independent "school" in the admission process. A student can apply for both regular and high priced seats to increase her chance of getting into that school.

After students enter high school, they spend the first year studying the same courses. At the end of their freshmen year, they have to decide whether to pursue the "science" or "arts" path then are sorted into respective classrooms and curriculum. If a student chooses to concentrate on science subjects, then in the CET, she will be tested on Chinese language, English language, Mathematics for science students, Physics, Chemistry and Biology. For arts students, she will take the CET on Chinese language, English language, Mathematics for arts students, Political science, History and Geography. The Chinese language and English language tests are identical for both arts and science students[7]. Like in the HET, the total score of one's CET is the sole determinant of college admission for most students. A science track student still takes classes of History, Political science and Geography after they made their track choice

---

[5] For details of the admission process, please refer to our other paper, which take advantage of the admission cutoff scores (Hoekstra, Mouganie, & Wang, 2016).
[6] This fee ranges from 20,000 to 40,000 RMB (about $3,000 - $6,000 USD) depending on the school and is public information.
[7] In the first three years of our sample, the CET takes the form of 3+X+S/A where a student will take the test on Chinese language, Mathematics (for science or arts), English language, one science or arts subject of her choice and comprehensive science or arts test. A science student can choose any of the three science subjects (Physics, Chemistry or Biology) as her "X" subject and an arts student chooses her "X" subject from Political science, History and Geography.

and vice versa. All students need to pass an assessment exam of all subjects at the end of their second year in high school regardless of track choice in order to graduate, which is a necessary condition to take the CET and go to college. The assessment test is generally very basic and nearly all students pass. The high school graduation rate in our sample is 99.66% as of 2012. However, the CET tests are meant to be more challenging so as to properly differentiate abilities. So it is general practice that schools prepare students as much as possible on their chosen subjects, and just enough to pass the assessment exam on the subjects that are not going to be on their CET. This disparity in training makes it very unlikely that a student switch from science to arts or the other way around after some time.

In college admission, colleges get to decide whether they want to admit science students or arts students or both for each of their majors. This information is given to students before they take the CET and is fairly consistent over time. In 2007, more than 60% of the college seats available for the sample province are for science students and around 40% are for arts students.

There are many reasons why a student may choose the science or arts path besides interests in the subject itself. The prospect of going to college is one of the most important considerations in this decision process. For example, the cutoff scores for science was significantly lower than arts in 2007 in our sample and we observe a surge in the proportion of students choosing science for both sexes in the class of 2009[8]. One's mathematics ability also plays a crucial role in that the CET mathematics test for arts student is easier than the version for science students.

IV.2.2 Model

We estimate the peer effect of high performing females following the reduced-form equation:

$$y_{ist} = \alpha_s + \beta_t + \boldsymbol{X}'_{ist}\lambda_1 + \boldsymbol{S}'_{st}\lambda_2 + \pi P_{ist} + \varepsilon_{ist} \qquad \text{(IV.1)}$$

where i denotes individuals, s denotes schools, and t denotes time. y_{ist} equals to 1 if the individual student i in school s at time t chooses the science path and 0 if she chooses the arts path. α_s is a school effect and β_t is a time effect. $\boldsymbol{X}_{ist}$ is a vector of student's covariates which includes gender, her entering scores by subject and indicators for private high school and high price payer. $\boldsymbol{S}_{st}$ is a vector of school characteristics of school s at time t, including average entering scores by subject of the cohort. P_{ist} is the proportion of high performing female students in school s at time t excluding student i herself. ε_{ist} is the error term, composed of school, time and individual specific random elements.

Our outcome is whether student i chooses take the CET in science, with treatment variable being the high performing female student ratio in the same school of

[8] The CET takes place every June and the admission cutoff information is made public in July. Since high school students choose their concentration at the end of their first year, Class of 2009 made this decision in the summer of 2007, possibly incorporating the 2007 CET information.

the same year as student i excluding herself, defined as the percentage of female students with HET math score at the top 20% level[9].

The proportion of high performing females is correlated with school quality. Higher achieving girls are more likely admitted by better high schools. Female students' major choices may also be affected by some of the school characteristics such as teacher quality and teacher gender ratio. To solve these endogeneity problems, we use within school variations in proportion of high performing female students across adjacent cohorts. This way we compare the science path choices of students from adjacent cohorts who have similar characteristics and face the same school environment, and the only variation is the change in the percentage of high achieving female students due to purely random factors. The coefficient π captures the effect of having more high performing female students on choosing science as major.

IV.3 Data

IV.3.1 Data Description

We have official administrative data from a large metropolitan area in southern China for four consecutive cohorts (students graduating high school from 2007 to 2010) of 12th grade students. Each record contains an individual identifier, a school identifier, the student's HET and CET scores by subject, some demographic information on the student: gender, minority status, parental occupation. It is possible that a student takes

[9] We will show that the results are robust using different specifications.

the CET more than once but we only observe the first attempt[10]. To uphold our assumption of fixed school quality, we drop schools that don't have all four years of enrollment due to either opening or closure within our sample period[11]. We further restrict the sample to the class of 2007, 2009 and 2010, dropping the class of 2008 when one district took a different HET test from the rest of the city rendering the test scores incomparable across the city.

IV.3.2 Summary Statistics

Table IV.1 Summary Statistics by Cohort and Gender

	2007		2009		2010	
	Female	Male	Female	Male	Female	Male
Percentage in cohort	52.7%	47.3%	52.5%	47.5%	52.1%	47.9%
Percentage choose science	27.6%	64.8%	41.4%	73.5%	33.2%	70.3%
Percentage in private school	<1%	<1%	2.4%	3.7%	2.5%	3.5%
Percentage entered paying high price	9.8%	13.2%	8.6%	12.5%	7.0%	9.4%
HET Chinese language	0.26	-0.29	0.21	-0.23	0.27	-0.29
HET Mathematics	-0.11	0.13	-0.09	0.10	-0.13	0.14
HET English language	0.20	-0.22	0.21	-0.23	0.22	-0.24
HET Physics	-0.17	0.19	-0.14	0.16	-0.15	0.16
HET Political science	0.18	-0.20	0.15	-0.17	0.17	-0.19
HET Chemistry	-0.14	0.16	-0.12	0.13	-0.13	0.14
High performing peer	8.3%	11.2%	9.1%	11.9%	9.1%	12.1%
CET Total Score (Arts)	0.12	-0.28	0.15	-0.36	0.13	-0.31
CET Total Score (Science)	-0.10	0.20	-0.8	0.14	-0.11	0.22
Number of students	22005	19726	24065	21744	24946	22972
Total number of students	41731		45809		41918	
Total percentage choose science	45.2%		56.6%		51.0%	

Notes: The table reports the means of the variables on the left. The HET and CET scores are standardized by subtracting the mean then divided by the standard deviation of the cohort sample. High performing peer is defined as the proportion of students whose HET mathematics score is among the top 20% of her cohort.

[10] If she decides to take the CET again, she has to take the test with the next cohort a year later. It is also unlikely that taking the test multiple times should affect her path choice as switching to the other path means she will have only one year to prepare for the new subjects while her competitors have two.
[11] 25 schools, 6,330 observations were excluded from the sample. Including these observations does not change the main findings. Results available upon request.

The remaining sample contains 118 high schools and 135,458 observations for the class of 2007, 2009 and 2010. Average class size is 383. We do not have classroom information so our peer statistics are all at the school-cohort level[12]. As shown in Table IV.1, percentage female in each cohort is around 52%. Private school enrollment is less than 4% of the population and about 10% of the students pay a high price to get into high school. In the HET test, girls on average do better in subjects that emphasis on language, writing and recall of facts such as Chinese and English language and Political science. Boys on the other hand perform better in quantitative and analytical fields: Mathematics, Physics and Chemistry. Of the class of 2007, only 45% chose to take the CET in science. For seven years before 2007, the CET in the sample province takes the form of 4+X. The four stables are Chinese and English language and Mathematics, and one comprehensive test of all 6 subjects: Physics, Chemistry, Biology, Political science, History and Geography. These four tests are the same to every student. The "X" subject is chosen by students from one of the 6 subjects in the comprehensive test. The total score of CET is calculated by the sum of the standardized scores of each subject. To maximize one's total score, a student should choose the "X" subject of her best comparative advantage instead of absolute ability. The science subjects are perceived as the more difficult ones and students who choose science subjects are more likely the better students from higher ranked high schools. There has been consistent decrease in students choosing Physics, Chemistry or Biology and the authorities decided to reform

---

[12] The classroom level statistics may not be a good measurement as classroom assignment can be endogenous (Lavy & Schlosser, 2011).

the CET for the 2007 cohort. In the 2007 CET, students choosing Physics, Chemistry or Biology are called the Science track and the rest of students are the Arts track. The comprehensive test and Mathematics are also made into two different versions for science and arts students respectively. Colleges decide whether to enroll science or arts students or both for each major. For most cases, science students and arts students compete among themselves respectively. However the decision was not made until 2006 summer, only a year before the 2007 cohort takes the CET, when they have already spent one year preparing for their chosen field. It is quite unlikely that a student switch field only a year before taking the test and even less likely to switch from a science subject to an arts subject or vice versa. Hence only 45% of the 2007 cohort chose science. The number rose to 52.8% in the 2008 cohort.

The treatment of interest in this paper is the effect of the high performing peers on choosing the science track. We define the high performing as the ability to be within the top 20% in the HET mathematics test. Then we calculate the female/male high performing ratio for student i by computing the percentage of the high performing female/male students in individual i's year-school cohort excluding herself/himself. The mean of high performing female peer ratio is 8.7% and 11.8% for males. Later on we also show that using a different specification of high performing students have similar results.

The cohort-to-cohort variations in the proportion high performing female could be random within a school; one could still be concerned that students might respond to these unpredicted shocks. They might know on average how well female students do in a certain school, but it is unlikely that they can predict in advance of the cohort that enters the school in a particular year as students make their choice of high schools before they even take the test. Although theoretically they could transfer to another school after they are exposed to this information after they entered the school, it is highly unlikely in reality. In China, one cannot transfer from one public high school to another unless you move to another city, so the only options are dropping out or transfer to a private high school.

We formally define high performing female students as the girls whose HET mathematics scores are among the top 20% of the cohort. In our sample, girls represent 42.5% of the top 20% math scorer in 2007, and the numbers are 43.1% and 42.1% for 2009 and 2010. T-test results also show that girls' average performance in HET mathematics is statistically the same throughout the years.

IV.4 Results

School choice is mostly endogenous and the number of high performing female students is closely associated with school quality. To account for the endogeneity in school choice and school quality, we rely on idiosyncratic variations in the proportion of female students who are good at mathematics within the same school but in different cohorts. Using this school-year fixed effect model, we examine whether cohort-to-cohort changes in the likelihood of choosing the science path is associated with cohort-to-cohort changes in the high performing female ratio within the class, assuming students from adjacent cohorts have similar characteristics and face the same school environment and teacher quality.

The effect of the proportion of high performing female students on science track choice is reported in Table IV.2. Our sample includes 135457 students from three cohorts. About 52% of the students are female in all cohorts.

Table IV.2 Estimates of the Effect of Proportion High Performing Peer on Science Track Choices in High School

	(1)	(2)	(3)	(4)	(5)
Proportion High Performing Female	0.367*	0.231	0.184	0.179	0.291
	(0.163)	(0.175)	(0.185)	(0.173)	(0.174)
Proportion High Performing Male	0.194	0.305	0.0411	0.334*	0.188
	(0.157)	(0.162)	(0.160)	(0.155)	(0.156)
Female × High Performing Female	0.356**	0.356**	0.354**	0.479***	0.268**
	(0.119)	(0.119)	(0.119)	(0.109)	(0.0998)
Female × High Performing Male	-0.231	-0.229	-0.227	-0.760***	-0.413***
	(0.120)	(0.120)	(0.120)	(0.114)	(0.106)
Proportion Female Peer		0.207*	0.0900	0.00867	-0.0136
		(0.0923)	(0.0960)	(0.0942)	(0.0973)
Female	-0.361***	-0.362***	-0.362***	-0.160***	-0.167***
	(0.00718)	(0.00719)	(0.00718)	(0.00700)	(0.00681)
High School Fixed Effects	Yes	Yes	Yes	Yes	Yes
Year Fixed Effects	Yes	Yes	Yes	Yes	Yes
Peer Mean HET Scores			Yes	Yes	Yes
Individual HET Scores				Yes	Yes
Relative Ranking within School					Yes
Observations	135458	135458	135458	135457	132910
R2	0.143	0.143	0.144	0.268	0.281

Standard errors in parentheses

*** Significant at the 1 percent level.

** Significant at the 5 percent level.

* Significant at the 10 percent level.

Notes: These are the notes applicable to the table. The style is Tables Notes.

The columns show the Year-School fixed effects estimates of the gender-heterogeneous effects of high performing female students on science track choices. Different controls are included in different columns in order to assess how sensitive the estimates are to individual characteristics. Column 1 presents the estimates with only the treatment variables and student gender as control. With this specification, the results show that the high performing female ratio will increase a male student's chance to choose science at the 95% confidence level. For female students, this effect is even bigger and more precise. Having more male classmates who are good at mathematics, however, have a smaller and statistically insignificant from zero effect on boys and potentially negative effect on girls. This means, while boys are more likely to choose science when they are surrounded with students better at math regardless of gender, girls are more encourage to take the science path when there are more high performing classmates of the same sex. Also being female alone significantly decreases the probability of choosing science. It is worth noting that the estimates for year fixed effects are sizeable and significant. Everything else being equal, the class of 2009 is 10% more likely to choose science than the class of 2007. This is due to the fact that the 2009 cohort expected a chance of going to college for science students[13].

Then we add proportion of female students as another control in column 2. Literature has shown that having more girls in a classroom can raise academic outcomes for both sexes through lower level of classroom disruption and violence (Lavy & Schlosser, 2011). We want to see if this could also affect the likelihood of girls choosing

[13] The 2009 cohort observes the 2007 cohort CET results at the end of their freshmen year, which is when they decide which path to take. In 2007, there were more students taking arts than science but college seats available for the province in science are 35% more than those in arts, resulting a higher cutoff score line (528 out of 750 points for four-year colleges) for arts students and a considerably lower one for science students (480 out of 750 points for four-year colleges). This information is made public by the media and thus affected the decision of the later cohorts. In 2008, the arts-science enrollment score gap shrank to only 6 points and the 2010 cohort was only 5% more likely to choose science than the 2007 cohort.

science subjects as well. Adding proportion of female into the equation does not seem to change the direction or precision of the coefficients from column 1 and it seems to increase students' likelihood to choose science at the 5% level. However this effect goes away as we add more individual characteristic controls in the following columns. In column 3, cohort controls are also included cohort controls such as the peer mean entering scores by subject and whether the high school is a private school. We see that adding these controls for cohort characteristics does not alter the signs of the estimated coefficients of the treatment variables.

The decision of which path to choose should also be correlated with the student's own relative strength as the ultimate goal is to score as high as possible in the CET so that she could go to a better college. Therefore we include the individual HET scores as measurement of their ability. As expected, students who are better at Math, Physics or Chemistry when entering high school are more likely to choose science as opposed to those who are better at Chinese language, English language or Politics. These individual controls seem to account for a large part of the female preference against the science track by decreasing the magnitude of the coefficient on the gender variable. This means girls' inclination away from science is partly because they score lower in HET Math, Physics and Chemistry on average. Girls could be discouraged by not doing well in science tests, or they decided to pursue the arts track where they have comparative advantage in. The directions and precision of the coefficients on our treatment variables are essentially unchanged as we incorporate cohort and individual features. For boys, the proportion of high performing peers of either sex has a moderately positive effect on whether they choose science. However for girls, a 10 percentage point increase in the proportion of girls in her cohort who are top 20% math scorer at math will raise her likelihood of choosing science by 4.8%, while a 10 percentage point increase of male top math scorers will reduce this likelihood by 7.6%. In column 5, we also include the within school rankings of their entering scores, taking in consideration the possibility that one's confident in learning may be correlated with their relative place in her class. The main results remains unchanged.

As previously discussed, one possible explanation for the gender gap in science and engineering field is lack of female role models in these fields. The high performing female peers might serve as the female role models for the rest of the girls in the class. It could also a form of affirmation which is consistent with the finding that value affirmation can raise women's performance in physics class but not effective for men (Miyake et al., 2010). Seeing these girls doing well in quantitative and analytical fields, her fellow female classmates could feel inspired and encouraged in learning these subjects and believe she can do well too.

Table IV.3 Results for Different Specifications of High Performing Peers

Variables	(1) 15%	(2) 20%	(3) 25%	(4) 30%	(5) 35%
Proportion High Performing Female	0.212	0.179	0.0966	0.0862	0.0828
	(0.200)	(0.173)	(0.143)	(0.121)	(0.119)
Proportion High Performing Male	0.315	0.334*	0.364*	0.278*	0.196
	(0.201)	(0.155)	(0.141)	(0.128)	(0.124)
Female × High Performing Female	0.662***	0.479***	0.315***	0.208**	0.0897
	(0.171)	(0.109)	(0.0861)	(0.0691)	(0.0613)
Female × High Performing Male	-0.913***	-0.760***	-0.618***	-0.535***	-0.422***
	(0.169)	(0.114)	(0.0975)	(0.0826)	(0.0759)
Proportion Female Peer	0.0407	0.00867	0.0501	0.0322	0.0307
	(0.0902)	(0.0942)	(0.0979)	(0.0958)	(0.0992)

Table IV.3 Continued

Variables	(1) 15%	(2) 20%	(3) 25%	(4) 30%	(5) 35%
Female	-0.173***	-0.160***	-0.151***	-0.140***	-0.133***
	(0.00647)	(0.00700)	(0.00761)	(0.00808)	(0.00855)
High School Fixed Effects	Yes	Yes	Yes	Yes	Yes
Year Fixed Effects	Yes	Yes	Yes	Yes	Yes
Peer Mean HET Scores	Yes	Yes	Yes	Yes	Yes
Individual HET Scores	Yes	Yes	Yes	Yes	Yes
Relative Ranking within School					
Observations	135458	135458	135458	135457	132910
R2	0.143	0.143	0.144	0.268	0.281

Standard errors in parentheses

*** Significant at the 1 percent level.

** Significant at the 5 percent level.

* Significant at the 10 percent level.

Notes: The specification used here follows column 4 of Table IV.2, with high performing female/male defined as the top 15%, 20%, 25%, 30% and 35% respectively.

Table IV.3 presents the robustness test with alternative definition of high performing peers. Previously we have defined the “high performing peer” as the students who are top 20% in HET math. Here we also check for other specifications: top 15%, top 25%, top 30%, and top 35%[14]. The main findings remain the same and the heterogeneous treatment effects are less significant as the high performing peer becomes more diluted which is consistent with the finding that gender ratio does not seem to affect how students make the decision of whether to choose science or arts. The estimates on controls for cohort and individual characteristics are consistent both in signs and magnitude.

[14] For top 5% and top 10% specifications, there are many schools have none of these top students.

IV.5 Conclusion

In this paper, we empirically estimate the heterogeneous treatment effect of high performing peers on high school students' choice of the science track. The unique data and Chinese setting allow us to track the choice of STEM fields as early as high school. The evidence provided in this paper suggests that the effects of high performing peers on boys are moderate to none. For girls however, having more high performing female peers can increase the likelihood for them to choose science while having more male peers who are good at math could potentially harm their chances to take the science track. One explanation is that girls might perceive her high performing female classmates as role models and an affirmation that they can do as well as them in science subjects. Unfortunately we do not have information on teachers, as teachers are also a form of role models and a source of value affirmation. It would be interesting to find out whether having more female science teachers could encourage girls to take the science track.

The individual HET scores account for a large part of the gender gap in science fields. Therefore in order to encourage girls to enter the science fields, essentially is to better prepare girls academically for science subjects by stimulating their interests or creating better early experience with science (Clark Blickenstaff, 2005). After accounting for the observable measurement, there still exist unobservable factors that cause girls to be 15% less likely to take the science track than boys. In the context of China, there is still negative gender stereotype endorsement and the perception that men are better than women in science and engineering. A 2011 survey on Chinese college students majoring in science and engineering shows that 10% of the female students agree or somewhat agree with the notion of "men are born to be better than women" and the number for male students is 23.1% (L.Zhang & H.Zhen, 2011). It is also worth noting that students, both male and females respond to immediate incentives of learning science. In our case it is the chance of going to college and it might extend to job opportunities and wage rates. However unlike going to college, the experience of looking for and being in a job can be different for women and men. More than 44% of the female STEM students in China reported "sex discrimination" in the job market and 53.8% of them would have chosen a major with less science component (L.Zhang & H.Zhen, 2011). Our results provide some suggestions on how to encourage girls to choose science in school, but more support on the job market are needed to make them stay in this industry.

CHAPTER V

CONCLUSIONS

This dissertation examines issues related to high school student test performance and track choice. Using student-level data from a large city in China, I explore heterogeneity in the effect of relative ranking, peer and teacher quality on student outcomes. My results from the paper show that attending better schools are generally beneficial in providing better peer and teacher quality. However, students at the bottom do suffer from a negative psychological effect that may hurt their test score outcomes to some extent. As of the decision on choosing the arts or science track, girls benefit from having more female peers who are strong in math. On the other hand, boys' choices don't seem to be affected by peers and the presence of high performing boys does not help girls become more interested in studying science.

REFERENCES

Andrews, R.J., Li, J., Lovenheim, M.F., (2012). Quantile Treatment Effects of College Quality on Earnings: Evidence from Administrative Data in Texas. *NBER Working Paper,* 18068.

Abdulkadiroglu, A., Angrist, J., Pathak, P., (2014). The Elite Illusion: Achievement effects at Boston and New York Exam Schools. *Econometrica*, 82(1): 137-196.

Angrist, J., Lavy, V., (1999). Using Maimonides' Rule to Estimate the Effect of Class Size on Scholastic Achievement. *The Quarterly Journal of Economics,* 114(2): 533-575.

Beede, D. N., Julian, T. A., Langdon, D., McKittrick, G., Khan, B., & Doms, M. E. (2011). Women in STEM: A gender gap to innovation. *Economics and Statistics Administration Issue Brief* (04-11).

Berkowitz, D., Hoekstra M., (2011). Does high school quality matter? Evidence from admis- sions data. *Economics of Education Review*, 30(2011): 280-288.

Calonico, S., Cattaneo, M,D., Titiunik, R., (2014). Robust Nonparametric Confidence Intervals for Regression-Discontinuity Designs. *Econometrica*, 82(6): 2295-2326.

Canaan, S., Mouganie, P., (2015). Returns to Education Quality for Low Skilled Students: Evidence from a Discontinuity. *SSRN Working Paper,* available at SSRN: http://ssrn.com/abstract=2518067.

Chetty R., Friedman, J.N., Rockoff, J.E., (2014). Measuring the Impacts of Teachers I: Evaluating Bias in Teacher Value-Added Estimates. *American Economic Review*, 104(9): 2593-2632.

Clark, D., (2010). Selective Schools and Academic Achievement. *B.E Journal of Economics Analysis and Policy: Advances in Economic Analysis and Policy*, 10(1).

Clark, D., Del Bono, E., (2014). The Long-Run Effects of Attending an Elite School: Evidence from the UK. *IZA Discussion Paper*, No. 8617 November 2014.

Clark Blickenstaff, J. (2005). Women and science careers: leaky pipeline or gender filter? *Gender and Education,* 17(4): 369-386.

Collins, C., Gan, L., (2013). Does Sorting Students Improve Scores? An Analysis of Class Composition. *NBER Working Paper*.

Davis, J, (1966). The campus as a frogpond: An application of the theory of relative deprivation to career decisions for college men. *American Journal of Sociology,* (72), 17-31.

Ding, W., Lehrer, S, F., (2007). Do Peers Affect Student Achievement in China's Secondary Schools? *Review of Economics and Statistics,* 89(2): 300-312.

Dobbie W., Fryer, R.G., (2014). The Impact of Attending a School with High-Achieving Peers: Evidence from the New York City Exam Schools. *American Economics Journal: Applied Economics*, 6(3): 58-75.

Duflo, E., Dupas, P., Kremer, M., (2011) Peer effects, teacher incentives, and the impact of tracking: Evidence from a randomized evaluation in Kenya. *American Economic Review*, 101(5).

Giles, J., Park, A., Wang, W., (2015). The Great Proletarian Cultural Revolution, Disruptions to Education, and the Returns to Schooling in Urban China. *IZA Discussion Paper,* No. 8930, March 2015.

Figlio, D., Page, M., (2002). School Choice and the Distributional Effects of Ability Tracking: Does Separation Increase Inequality? *Journal of Urban Economics*, 51: 497-514.

Hannum, E., Park, A., (2001). Do Teachers Affect Learning in Developing Countries?: Evidence from Matched Student-Teacher Data from China. *Conference Rethinking Social Science Research on the Developing World in the 21st Century*.

Hoekstra, M., (2009). The Effect of Attending the Flagship State University on Earnings: A Discontinuity-Based Approach. *The Review of Economics and Statistics,* 91(4): 717-24.

Hoekstra, M., Mouganie, P., Wang, Y. (2016). Peer Quality and the Academic Benefits to Attending Better Schools. *Working paper*.

Hyde, J. S., Fennema, E., Lamon, S. J. (1990). Gender differences in mathematics performance: a meta-analysis. *Psychological bulletin,* 107(2): 139.

Hyde, J. S., Lindberg, S. M., Linn, M. C., Ellis, A. B., & Williams, C. C. (2008). Gender similarities characterize math performance. *Science,* 321(5888): 494-495.

Imbens, G.W., Lemieux, T., (2008). Regression discontinuity designs: A guide to practice. *Journal of Econometrics,* 142(2): 615-35.

Jackson, C.K, (2010). Do Students Benefit From Attending Better Schools? Evidence From Rule-based Student Assignments in Trinidad and Tobago. *Economic Journal*, 120(549): 1399-1429.

Jackson, C.K, (2013). Can higher-achieving peers explain the benefits to attending selective schools? Evidence from Trinidad and Tobago. *Journal of Public Economics*, 108: 63-77.

Kane, T.J., Rockoff, J.E., Staiger, O., (2008). What Does Certification Tell Us about Teacher Effectiveness? Evidence from New York City. *Economics of Education Review,* 27(6): 615-31.

Krueger, A.B., (2003). Economic Considerations and Class Size. *Economic Journal,* 113(485): F34-F63.

Lai, F., Sadoulet, E., de Janvry, A., (2011). The Contributions of School Quality and Teacher Qualifications to Student Performance: Evidence from a Natural Experiment in Beijing Middle Schools. *Journal of Human Resources,* 46: 123-153.

Lavy, V., Paserman, M. D., Schlosser, A. (2012). Inside the Black Box of Ability Peer Effects: Evidence from Variation in the Proportion of Low Achievers in the Classroom. *The Economic Journal*, 122(559): 208-237.

Lavy, V., Schlosser, A. (2011). Mechanisms and impacts of gender peer effects at school. *American Economic Journal: Applied Economics*, 1-33.

Lee, D.S, (2009). Training, Wages, and Sample Selection: Estimating Sharp Bounds on Treatment Effects. *Review of Economic Studies*, 76(3): 1071-1102.

Lee, D.S., Card, D. (2008). Regression discontinuity inference with specification error. *Journal of Econometrics*, 142(2): 655-674.

Lee, D.S., Lemieux, T., (2010). Regression discontinuity designs in economics. *Journal of Economic Literature,* 48(2): 281-355.

Lefren, L., (2004). Educational Peer Effects and the Chicago Public Schools. *Journal of Urban Economics*, 56(2): 169-191.

Li, H., Liu, P. W., Ma, N., Zhang, J., (2005). Does Education Pay in Urban China? Estimating Returns to Education Using Twins. *Discussion Papers* 00015, Chinese University of Hong Kong, Department of Economics.

Lucas, A.M., Isaac M.M., (2014). Effects of School Quality on Student Achievement: Dis- continuity Evidence from Kenya. *American Economic Journal: Applied Economics,* 6(3): 234-63.

Marsh, H.W., (1984). Self-concept, social comparison, and ability grouping: A reply to Kulik and Kulik. *American Educational Research Journal*, 21(4): 799-806.

Miyake, A., Kost-Smith, L. E., Finkelstein, N. D., Pollock, S. J., Cohen, G. L., & Ito, T. A. (2010). Reducing the gender achievement gap in college science: a classroom study of values affirmation. *Science,* 330(6008): 1234-1237.

National Science Foundation (2014). Women, Minorities, and Persons with Disabilities in Science and Engineering. *National Science Foundation.*

Park, A., Xinzheng, S., Xuenhi, An., Hsieh C., (2015). Magnet High Schools and Academic Performance in China: A Regression Discontinuity Design. Forthcoming in *Journal of Comparative Economics*.

Pop-Eleches, C., Urquiola M., (2013). Going to a Better School: Effects and Behavioral Responses. *American Economic Review,* 103(4): 1289-1324.

Rivkin, S.G., Hanushek, E.A., & Kain, J.F., (2005). Teachers, Schools and Academic Achievement. *Econometrica,* 73(2): 417-58

Rockoff, J.E., (2004). The Impact of Individual Teachers on Student Achievement: Evidence from Panel Data. *American Economic Review: Papers and Proceedings,* 94(2): 247-52

Saavedra, J., (2009). The Learning and Early Labor Market Effects of College Quality: A Regression Discontinuity Analysis. *Rand Corporation Working Paper*.

Sacerdote, B., 2001. Peer Effects with Random Assignment: Results for Dartmouth Roommates. *Quarterly Journal of Economics,* 116(2): 681-704.

Seaton et al., 2010. Big-Fish-Little-Pond Effect: Generalizability and Moderation-Two sides of the same coin. *American Educational Research Journal*, 47(2): 390-433.

Turner, S. E., Bowen, W. G. (1999). Choice of major: The changing (unchanging) gender gap. *Industrial & Labor Relations Review,* 52(2), 289-313.

U.S. News & World Report (2015). Raytheon STEM Index Shows Gender and Racial Gaps Widening in STEM Fields. *U.S. News & World Report*. June 29, 2015

Zafar, B. (2013). College major choice and the gender gap. *Journal of Human Resources,* 48(3): 545-595.

Zimmerman, S., (2014). The Returns to College Admission for Academically Marginal Students. *Journal of Labor Economics,* 32(4): 711-754.

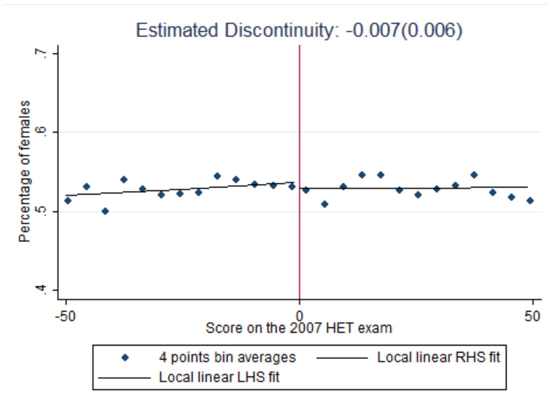
Zhang, H., (2014). The Mirage of Elite Schools: Evidence from Lottery-based School Admissions in China. *Unpublished Paper*, Chinese University of Hong Kong.

Zhang, L., Zhen, H., (2011). A Study on the Dilemma The Female University Students in Science and Technology Encountered during Their Learning. *Tsinghua Journal of Education,* 32(5).

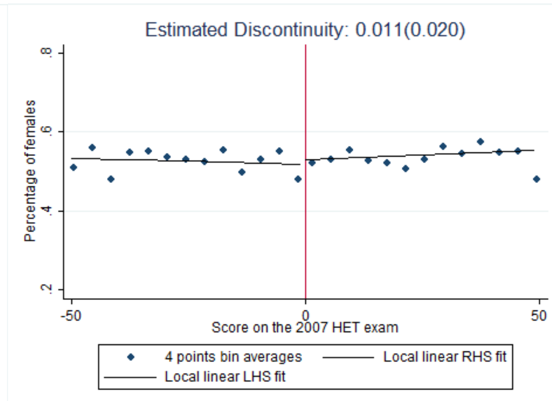
APPENDIX A FIGURES

Figure A 1 Smoothness of Baseline Covariates

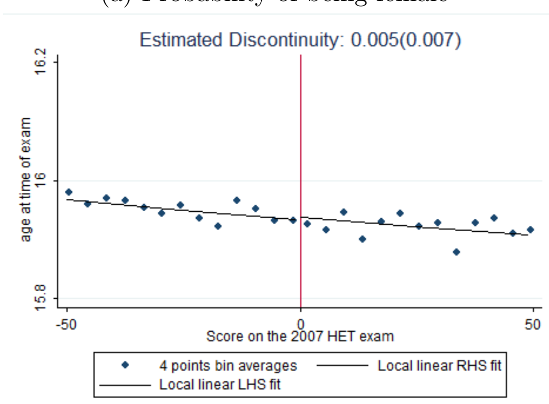
Attending Better Schools (All Cutoffs) | Attending a Tier I school



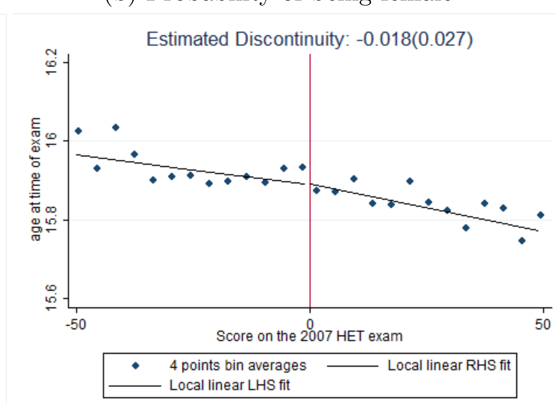
(a) Probability of being female



(b) Probability of being female



(c) Age at time of high school exam

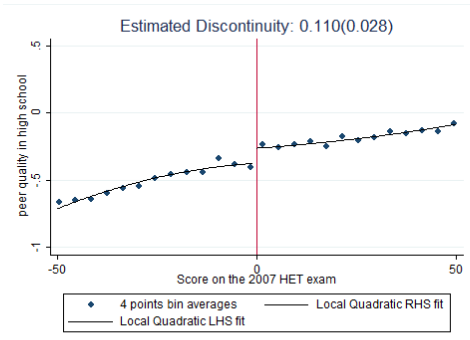


(d) Age at time of high school exam

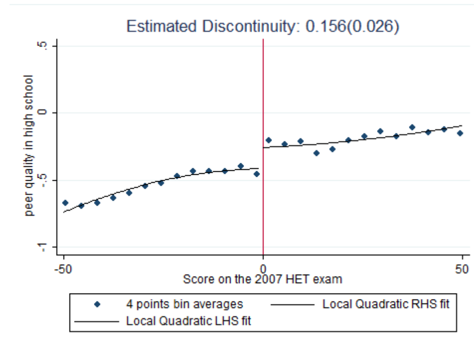
Notes: Sample includes students who took the high school entrance exam in the year 2007.

Figure A 2 Local Polynomial "Stacked RD" Estimates for Attending Better Schools, by Gender

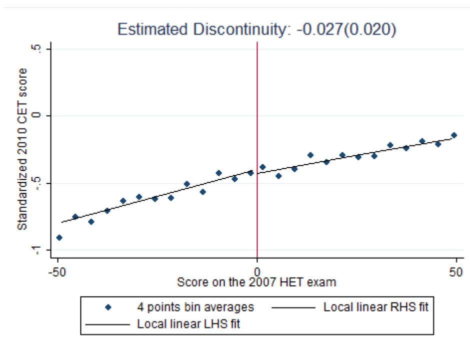
<u>Males</u> <u>Females</u>



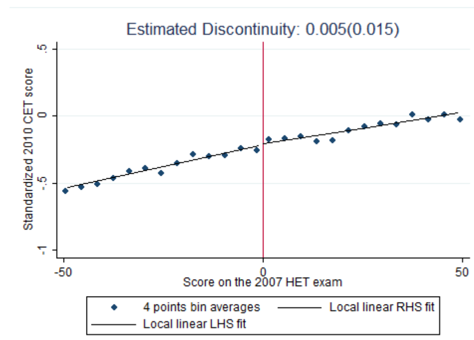
(a) Peer quality in high school



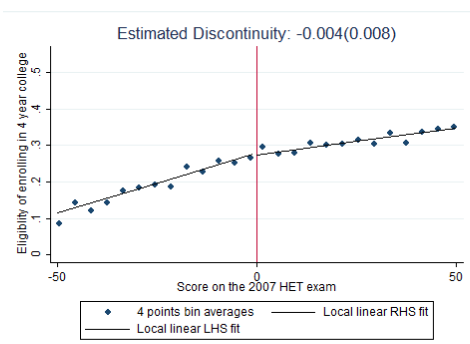
(b) Peer quality in high school



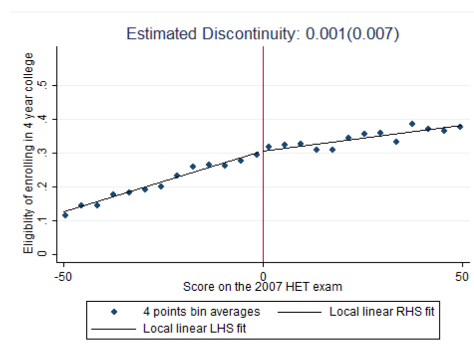
(c) CET exam scores



(d) CET exam scores



(e) Likelihood of enrolling in 4-year college



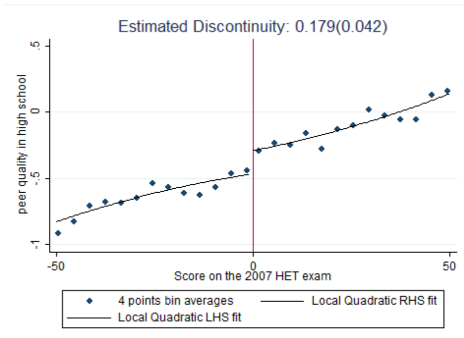
(f) Likelihood of enrolling in 4-year college

Notes: Sample includes students who took the high school entrance exam in the year 2007. Due to repeated observations, standard errors clustered at individual level.

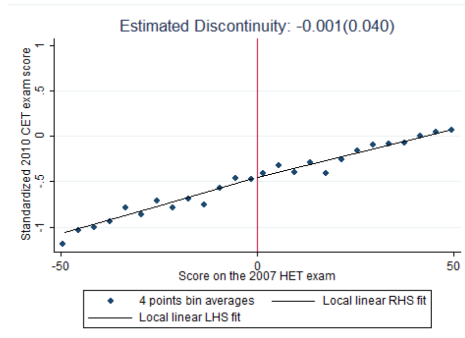
Figure A 3 Local Polynomial Peer Quality and Test Score Estimates for Other Cutoffs

High School Peer Quality

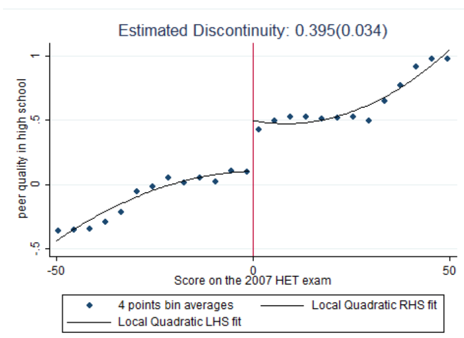
Standardized CET scores



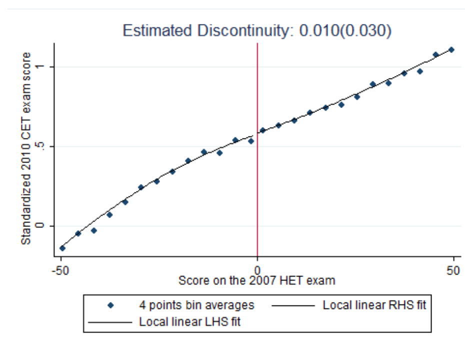
(a) Tier II school cutoff



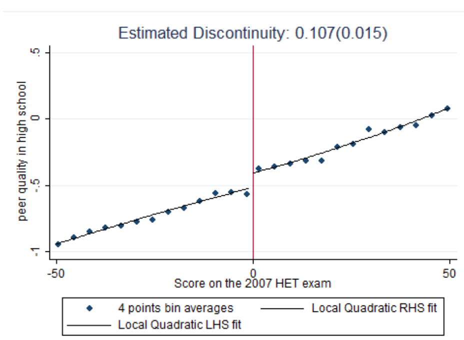
(b) Tier II school cutoff



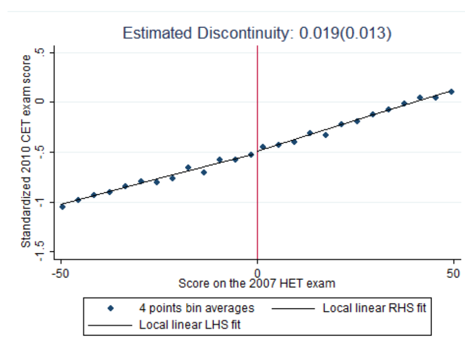
(c) Within Tier I school cutoffs



(d) Within Tier I school cutoffs



(e) Within Tier II school cutoffs

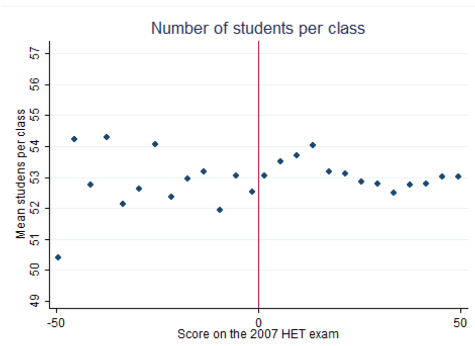


(f) Within Tier II school cutoffs

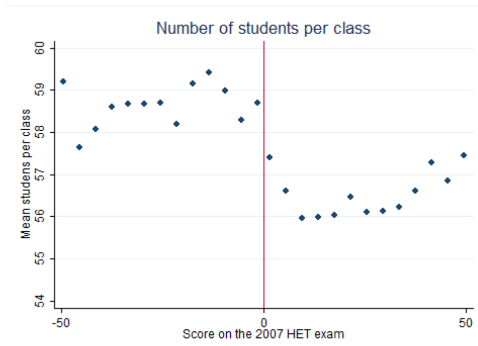
Notes: Sample includes students who took the HET exam in the year 2007. Standard errors clustered at the score level.

Figure A 4 Class Size and Teacher Quality for Other Cutoffs

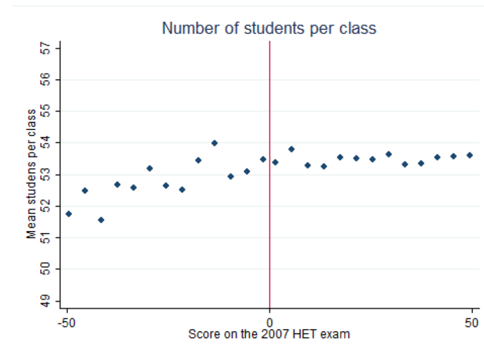
Tier II cutoff | Within Tier I cutoffs | Within Tier II cutoffs



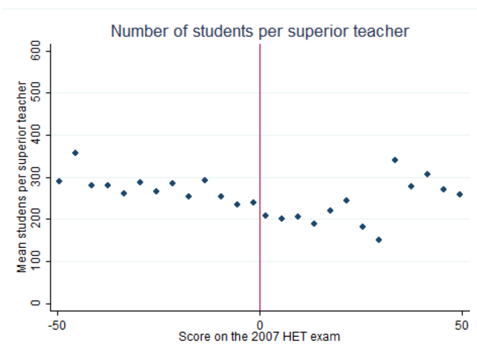
(a) Number of students per class



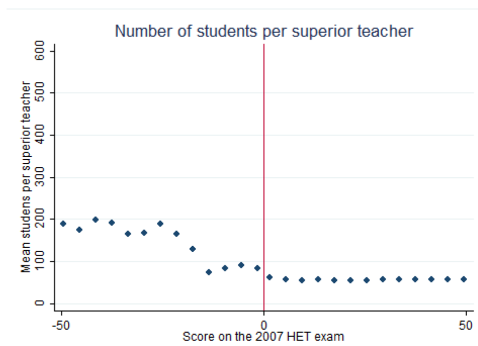
(b) Number of students per class



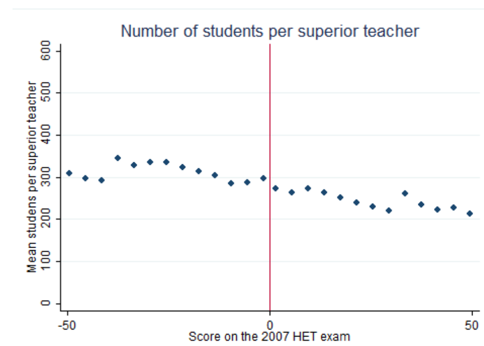
(c) Number of students per class



(d) Number of students per superior teacher



(e) Number of students per superior teacher



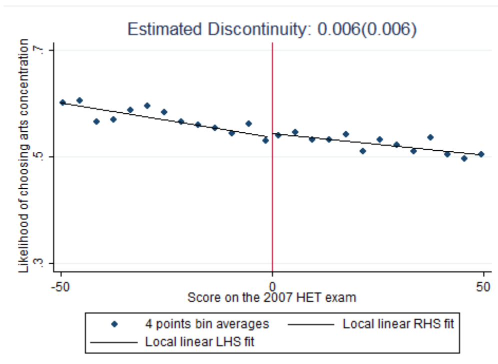
(f) Number of students per superior teacher

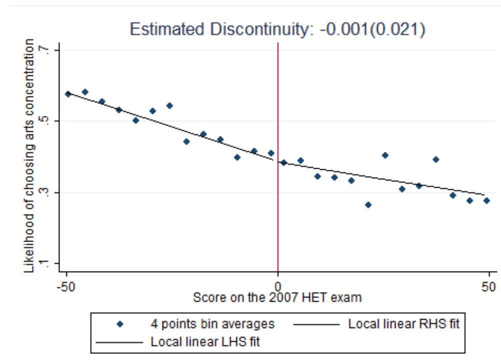
Notes: Samples based off of school level data.

Figure A 5 Threats to Interpretation

Attending Better Schools

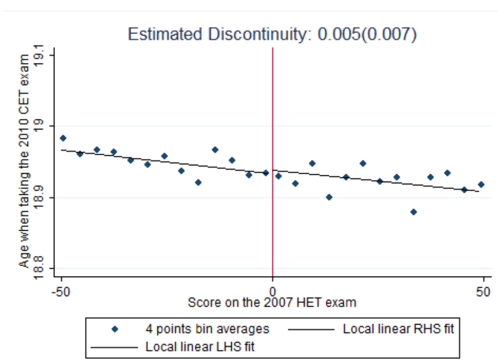
Attending a Tier I school

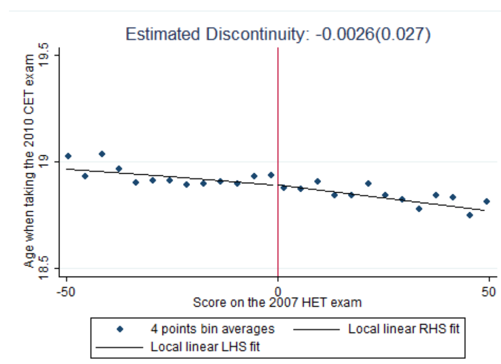




(a) Likelihood of majoring in arts versus sciences in High school.

(b) Likelihood of majoring in arts versus sciences in High school.





(c) Exact age when taking the 2010 CET exam.

(d) Exact age when taking the 2010 CET exam.

Notes: Sample includes students who took the CET exam in the year 2010.

APPENDIX B TABLES

Table A 1 Local Polynomial RD Estimates for Baseline Covariates

Table A1: Local polynomial RD estimates for baseline covariates

Bandwidth	2.5 CCT	2 CCT	1.5 CCT	1.25 CCT	1 CCT	0.75 CCT
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: (Going to a better school)						
Predicted HET score	−.006	−.005	−.004	−.001	−.002	.005
	(.004)	(.005)	(.006)	(.006)	(.007)	(.007)
Likelihood of being a female	−.004	−.006	−.007	−.007	−.008	−.001
	(.006)	(.007)	(.008)	(.008)	(.009)	(.010)
Age when taking HET entrance exam	.003	.005	.002	.006	.006	−.004
	(.008)	(.009)	(.010)	(.011)	(.013)	(.014)
Observations for predicted score	75977	62474	47880	40676	32499	24931
Panel B: (Going to a top-school)						
Predicted HET score	−.014	−.009	.005	.021	.007	.019
	(.014)	(.015)	(.017)	(.018)	(.020)	(.021)
Likelihood of being a female	.011	.017	.019	.029	.049**	.049*
	(.019)	(.020)	(.021)	(.022)	(.023)	(.027)
Age when taking HET entrance exam	−.011	−.012	−.003	−.006	−.024	−.051
	(.021)	(.022)	(.027)	(.029)	(.032)	(.036)
Observations for predicted score	6680	5578	4278	3642	2886	2237
Score Polynomial	One	One	One	One	One	One

Notes: Sample includes students who took the high school entrance exam in the year 2007 with known high school cutoffs. Predicted HET score based on the following controls: sex, gender, district fixed effects, middle school fixed effect. Optimal BW = 34 for predicted score, 34 for probability of being a female, 40 for age when taking HET exam (Going to a better school).
Optimal BW = 18 for predicted score, 26 for probability of being a female, 34 for age when taking HET exam (Going to a top school). *** p <0.01 ** p <0.05 * p <0.1

Table A 2 "Stacked Rd" Estimates for All Cutoffs Except the Tier I Cutoff

Bandwidth	18 points	16 points	14 points
	(1)	(2)	(3)
Panel A:			
Discontinuity in high school peer quality	.163***	.158***	.180***
	(.026)	(.028)	(.030)
With Controls	.166***	.161***	.183***
	(.02)	(.03)	(.03)
Panel B:			
Discontinuity in CET exam scores	.020	−.002	.014
	(.032)	(.035)	(.037)
With Controls	.028	.007	.018
	(.03)	(.03)	(.04)
Score Polynomial	One	One	One
Observations	14,624	13,055	11,432

Notes: Sample includes students who took the college entrance exam in the year 2007. Controls include: Age, gender, district fixed effects and middle school fixed effects. The maximum possible bandwidth under this identification strategy is 18 test score points. Since we observe individuals with multiple cutoffs, we cluster at the student ID level. *** p <0.01 ** p <0.05 * p <0.1

Table A 3 Robustness Check—Adding Third District to the Sample

Treatment effect	Original Sample	Add district 183 (5 Tier I schools)	Add district 183 (1 Tier I school)
Panel A: Going to a better school			
High school peer quality	0.170***	0.180***	0.180***
	(0.021)	(0.017)	(0.017)
College exam scores	-0.016	-0.014	-0.014
	(.014)	(0.022)	(0.022)
Likelihood of enrolling in four year college	–.002	.002	.002
	(.006)	(.011)	(.011)
Observations	37,961	53,334	53,334
Panel B: Going to a top school			
First Stage	.632***	.477***	.661***
	(.036)	(.025)	(.035)
High school peer quality	0.303***	0.210***	0.350***
	(0.026)	(0.024)	(0.028)
College exam scores	0.094***	0.083***	0.078***
	(0.024)	(0.022)	(0.024)
Likelihood of enrolling in four year college	.071***	.075***	.054***
	(.027)	(.026)	(.021)
Observations	6,046	8,056	8,056

Notes: Sample includes students who took the high school entrance exam in the year 2007. All regressions include controls: gender, age, district fixed effects and junior high school fixed effects. For ease of comparison, all regressions use an equal bandwidth of 40 score points on either side of the cutoff. *** p <0.01 ** p <0.05 * p <0.1